PB89-195036

Interrelation of Experimental Exposure and Ambient Air Quality Data for Comparison of Ozone Exposure Indices and Estimating Agricultural Losses

NSI Technology Services Corp., Corvallis, OR

Prepared for:

Corvallis Environmental Research Lab., OR

Jun 89

U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA/600/3-89/047	**2.**	**3. RECIPIENT'S ACCESSION NO.** PB89 195036/AS
4. TITLE AND SUBTITLE Interrelation of Experimental Exposure and Ambient Air Quality Data for Comparison of Ozone Exposure Indices & Estimating Agricultural Losses		**5. REPORT DATE** June 1989
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) E.Henry Lee[1], D.T. Tingey[2] , and W.E. Hogsett[2]		**8. PERFORMING ORGANIZATION REPORT NO.**
9. PERFORMING ORGANIZATION NAME AND ADDRESS [1]NSI, Corvallis, 2 EPA, Corvallis		**10. PROGRAM ELEMENT NO.**
		11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS US Environmental Protection Agency Environmental Research Laboratory 200 SW 35th Street Corvallis, OR 97333		**13. TYPE OF REPORT AND PERIOD COVERED** Published Report
		14. SPONSORING AGENCY CODE EPA/600/2

15. SUPPLEMENTARY NOTES

1989.

16. ABSTRACT

This report evaluates ambient air quality indicators for ozone (O_3) and estimates the exposure levels associated with agricultural losses. In support of the U.S. EPA Office of Air Quality Planning and Standards (OAQPS) review and evaluation of a secondary National Ambient Air Quality Standard (NAAQS) for ozone, the Terrestrial Branch within the Corvallis Environmental Research Laboratory (CERL) investigated alternative exposure indices for characterizing the adverse effects of ozone on agriculture, using experimental data from the National Crop Loss Assessment Network (NCLAN) program, Relationships among the current secondary NAAQS standard (second highest 1-h daily maximum concentration, denoted HDM2), the 7-h seasonal mean (0900 to 1600 h) adopted by NCLAN to relate plant response to O_3 exposure (M7), and the indices that "best" describe the exposure-response relationship were studied using 83 non-urban site-years of ambient air quality data from the National Air Data Branch (NADB).

Scientific and technical information on ozone and its effects on vegetation, natural ecosystems, natural and artificial materials, and human welfare necessitates a secondary standard that adequately describes the temporal variations of exposure and their effects on plant production and correlates well with short- and long-term exposures. The identification of an appropriate ozone standard that protects against crop yield losses proceeds along two lines: 1) regression analysis of plant response data against various exposure indices to determine indices that best depict biological response, and 2) ambient air quality analysis of ambient O_3 data to identify indices that correlate well with various exposure patterns.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT Release to Public	**19. SECURITY CLASS** *(This Report)* Unclassified	**21. NO. OF PAGES** 114
	20. SECURITY CLASS *(This page)* Unclassified	**22. PRICE**

EPA Form 2220-1 (9-73)

PB89-195036

EPA/600/3-89/047
June 1989

Interrelation of Experimental Exposure and Ambient Air Quality Data for Comparison of Ozone Exposure Indices and Estimating Agricultural Losses

by

E. Henry Lee
NSI Technology Services Corp.
200 SW 35th Street
Corvallis, Oregon 97333

and

David T. Tingey & William E. Hogsett
U.S. Environmental Protection Agency
Corvallis Environmental Research Laboratory
Corvallis, Oregon 97333

Contract No. 68-C8-0006

Project Officer

David T. Tingey
U.S. Environmental Protection Agency
Corvallis Environmental Research Laboratory
Corvallis, Oregon 97333

For
ENVIRONMENTAL RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
CORVALLIS, OR 97333

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL INFORMATION SERVICE
SPRINGFIELD, VA. 22161

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

TABLE OF CONTENTS

Section Title Page

FIGURES. iv

TABLES ix

ACKNOWLEDGEMENT. xi

CHAPTER 1. EXECUTIVE SUMMARY 1-1

CHAPTER 2. COMPARISON OF OZONE EXPOSURE INDICES IN EXPOSURE-RESPONSE MODELING. 2-1

2.1. Introduction 2-1

2.2. Evaluation of Exposure Indices 2-2

2.3. Methods 2-4

2.3.1. Exposure Response Data. 2-5
2.3.2. Exposure Indices. 2-6
2.3.3. Weibull Exposure-Response Model 2-8
2.3.4. Regression Approach to Compare Exposure Indices. 2-9

2.4. Results 2-10

2.5. Discussion. 2-12

2.6. Conclusion. 2-13

CHAPTER 3. ESTIMATION OF EXPOSURE LEVELS ASSOCIATED WITH VARIOUS CROP-YIELD LOSSES 3-1

3.1. Introduction. 3-1

3.2. Methods 3-2

3.3. Results 3-3

3.4. Discussion 3-6

3.5. Conclusion 3-7

TABLE OF CONTENTS

Section Title Page

CHAPTER 4. RELATIONSHIPS AMONG EXPOSURE INDICES AT 80 RURAL AND AGRICULTURAL SITES. 4-1

4.1. Introduction 4-1

4.2. Method 4-2

4.2.1. Ambient Ozone Monitoring Data 4-2
4.2.2. Ozone Exposure Indices 4-2
4.2.3. Regression Analysis and Joint Distributions 4-3

4.3. Results 4-3

4.3.1. Empirical Distributions of Exposure Indices 4-3
4.3.2. Relationships Among Exposure Indices . . 4-5
4.3.3. Joint Distributions of Exposure Indices . 4-7

4.4. Discussion 4-8

4.5. Conclusion 4-8

REFERENCES R-1

APPENDIX A. ABBREVIATIONS, CODE NAMES, AND FORMULAE A-1

APPENDIX B. AMBIENT AIR QUALITY INDICATORS AND ASSOCIATED PREDICTED RELATIVE YIELD LOSSES FOR AGRICULTURAL/REMOTE SITES IN THE U.S. B-1

B.1. Introduction. B-1

B.2. Methods B-2

B.3. Results B-3

B.4. Discussion. B-5

B.5. Conclusion. B-6

FIGURE LEGENDS

No.	Title	Page
1.1.	Boxplots of percentiles and mean predicted relative yield losses, characterizing the range of crop losses across the 16 NCLAN studies using the Weibull model with (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07 (same as Figure 3.6 in Chapter 3)	1-5
2.1.	Two comparisons of different exposure indices using the same biological response data	2-19
2.2.	Description of the methods used to calculate the areas under hourly concentration curve for the cumulative indices, AOT08 and SUM08	2-20
2.3.	Weighting functions used to calculate the generalized phenologically-weighted cumulative impact index: sigmoid (A) and allometric (B) functions for weighting the hourly ozone concentrations; (C) gamma functions mimicking plant sensitivity over time (assumes a 70-day exposure), and (D) three intra-episodic weights mimicking pollutant uptake during an episode	2-21
2.4.	Allometric and sigmoid functions used to weight the hourly ozone concentrations	2-22
2.5.	Examples of mathematical functions used to simulate changes in plant sensitivity with time	2-23
3.1.	Illustration of the exposure level associated with a 20% relative yield loss using the exposure-response model for 1983 soybean (Amsoy) study of Argonne, IL with the 7-h seasonal mean (M7)	3-12
3.2.	Predicted relative yield losses (PYRLs) for 16 NCLAN studies using the Weibull model and HDM2 at (A) 0.06 ppm, (B) 0.08 ppm, (C) 0.10 ppm, and (D) 0.12 ppm	3-13

FIGURE LEGENDS

No.	Title	Page
3.3.	Predicted relative yield losses (PRYLs) for 16 NCLAN studies using the Weibull model and M7 at (A) 0.04 ppm, (B) 0.05 ppm, (C) 0.06 ppm, and (D) 0.07 ppm	3-14
3.4.	Predicted relative yield losses (PRYLs) for 16 NCLAN studies using the Weibull model and SUM06 at (A) 5 ppm-h, (B) 10 ppm-h, (C) 20 ppm-h, and (D) 40 ppm-h	3-15
3.5.	Predicted relative yield losses (PRYLs) for 16 NCLAN studies using the Weibull model and SUM07 at (A) 5 ppm-h, (B) 10 ppm-h, (C) 20 ppm-h, and (D) 40 ppm-h	3-16
3.6.	Boxplots of percentiles and mean predicted relative yield losses, characterizing the range of crop losses across the 16 NCLAN studies using the Weibull model with (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07	3-17
4.1.	Cumulative distribution functions (CDF) and frequency distributions of (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07 from 83 agricultural/remote SAROAD site-years of ambient data	4-16
4.2.	Scatterplot matrix of the four exposure indices (HDM2, SUM06, SUM07, and M7) from 83 agricultural/remote SAROAD site-years	4-17
4.3.	Joint distributions of co-occurrence of (A) M7 and HDM2, (B) SUM06 and HDM2, (C) SUM07 and HDM2, (D) SUM06 and M7, (E) SUM07 and M7, and (F) SUM06 and SUM07, for 83 agricultural/remote SAROAD site-years of data	4-18
B.1.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1980 soybean study at Argonne, IL, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-9

FIGURE LEGENDS

No.	Title	Page
B.2.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1983 soybean study (Amsoy) at Argonne, IL, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-10
B.3.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1983 soybean study (Corsoy) at Argonne, IL, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-11
B.4.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1981 soybean study at Raleigh, NC, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-12
B.5.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1982 soybean study Raleigh, NC, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-13
B.6.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1982 wheat study (Abe) at Argonne, IL, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-14
B.7.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1982 wheat study (Arthur) at Argonne, IL, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-15

FIGURE LEGENDS

No.	Title	Page
B.8.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1983 wheat study (Abe) at Argonne, IL, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-16
B.9.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1983 wheat study (Arthur) at Argonne, IL, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-17
B.10.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1982 wheat study at Ithaca, NY, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-18
B.11.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1983 wheat study at Ithaca, NC, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-19
B.12.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1982 sorghum study at Argonne, IL, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-20
B.13.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1981 corn study (Pag-397) at Argonne, IL, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-21

FIGURE LEGENDS

No.	Title	Page
B.14.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1981 corn study (Pioneer) at Argonne, IL, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 .rm65 agricultural site-years	B-22
B.15.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1981 cotton study at Shafter, CA, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-23
B.16.	Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1982 cotton study at Shafter, CA, using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years	B-24
B.17.	Scatterplot of the 50th percentile of predicted relative yield losses (PRYLs) using (A) M7, (B) SUM06, and (C) SUM07, for the 83 agricultural site-years sorted in ascending order of HDM2	B-25
B.18.	Scatterplot of the 75th percentile of predicted relative yield losses (PRYLs) using (A) M7, (B) SUM06, and (C) SUM07, at the 83 agricultural site-years sorted in ascending order of HDM2	B-26
B.19.	Scatterplot of the 85th percentile of predicted relative yield losses (PRYLs) using (A) M7, (B) SUM06, and (C) SUM07, at the 83 agricultural site-years sorted in ascending order of HDM2	B-27
B.20.	The 50th, 75th, and 85th percentile of predictedrelative yield losses for the 16 NCLAN cases using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07 calculated at 83 non-urban areas	B-28

TABLES

No.	Title	Page
2.1	A. Performance criteria for 25 general ozone exposure and 3 other types of indices (GPWCI, SIGMOID, and ALLOMETRIC). B. Classification of 25 general ozone exposure indices ranked by average score in groups of 100 (Weibull model) . . .	2-14
2.2.	Classification of the 589 ozone exposure indices ranked by average score in groups of 100, and classified according to the sigmoid and allometric functions (Weibull model)	2-15
2.3.	Classification of the 589 ozone exposure indices ranked by average score in groups of 100, and classified according to the phenological weighting gamma functions (Weibull model)	2-16
2.4.	Classification of the 589 generalized phenologically-weighted cumulative impact indices ranked by average score in groups of 100, and classified according to the intra-episodic weights (Weibull model)	2-17
2.5.	Average scores for various combinations of the sigmoid and phenological gamma functions for calculating the generalized phenologically-weighted cumulative impact indices (Weibull model)	2-18
3.1.	Range of ozone treatments and exposure durations used in the 12 NCLAN studies	3-8
3.2.	Values of the parameters of the Weibull exposure-response functions for each of the 16 NCLAN study cases	3-9
3.3.	Percentiles and mean predicted relative yield losses (PRYLs) associated with various levels of the four exposure indices, HDM2, M7, SUM06, and SUM07, for the 16 NCLAN cases	3-10
3.4.	Exposure levels associated with predicted relative yield losses (PRYLs) of 5% to 30% for the four exposure indices, HDM2, M7, SUM06, and SUM07, for the 16 NCLAN studies	3-11
4.1.	Rural agricultural and remote monitoring sites reporting ozone data from SAROAD	4-10

TABLES

No. Title Page

4.2. Proportion of the annual SUM06 and SUM07 values that occur within a 3-, 4-, or 5-month period from May to September 4-11

4.3 Cumulative distribution functions (CDFs) of the four exposure indices, HDM2, SUM06, SUM07, and M7 for the 83 SAROAD sites 4-12

4.4. Relationships among exposure indices, HDM2, SUM06, SUM07, and M7 4-13

4.5. Joint distributions of the four exposure indices HDM2, M7, SUM06, and SUM07 4-14

B.1. Predicted relative yield losses (PRYL) for the 83 rural and agricultural sites based on the Weibull exposure-response models for 16 NCLAN studies B-7

B.2. Interrelation of experimental and ambient air quality data to estimate the exposure levels associated with mean, 50th, 75th, and 85th percentiles of PRYLs of 10% and the proportion of sites affected B-8

ACKNOWLEDGEMENT

The authors acknowledge contribution of data from principal investigators of those NCLAN field sites used in this paper: Lance W. Kress (Argonne, IL), Robert J. Kohut (Ithaca, NY), Allen S. Heagle (Raleigh, NC), and Patrick J. Temple (Shafter, CA). The research described in this article has been partially funded by the Office of Air Quality Planning and Standards, Durham, NC, and conducted at the Environmental Protection Agency's Research Laboratory in Corvallis, OR. The work upon which this publication is based was performed pursuant to Contract No. 68-C8-0006 with the U.S. Environmental Protection Agency.

CHAPTER 1

EXECUTIVE SUMMARY

This report evaluates ambient air quality indicators for O_3 and estimates the exposure levels associated with agricultural losses. In support of the U.S. EPA Office of Air Quality Planning and Standards (OAQPS) review and evaluation of a secondary National Ambient Air Quality Standard (NAAQS) for ozone, the Terrestrial Branch within the Corvallis Environmental Research Laboratory (CERL) investigated alternative exposure indices for characterizing the adverse effects of ozone on agriculture, using experimental data from the National Crop Loss Assessment Network (NCLAN) program. Relationships among the current secondary NAAQS standard (second highest 1-h daily maximum concentration, denoted HDM2), the 7-h seasonal mean (0900 to 1600 h) adopted by NCLAN to relate plant response to O_3 exposure (M7), and indices that "best" describe the exposure-response relationship were studied using 83 non-urban site-years of ambient air quality data from the National Air Data Branch (NADB).

Scientific and technical information on ozone and its effects on vegetation, natural ecosystems, natural and artificial materials, and human welfare necessitates a secondary standard that adequately describes the temporal variations of exposure and their effects on plant production and correlates well with short- and long-term exposures. The identification of an appropriate ozone standard that protects against crop yield losses proceeds along two lines: (1) regression analysis of plant response data against various exposure indices to determine indices that best depict biological response, and (2) ambient air quality analysis of ambient O_3 data to identify indices that correlate well with various exposure patterns.

EVALUATION OF EXPOSURE INDICES

Crop yield data from 12 NCLAN studies, with multiple cultivars for a total of 17 cases, were regressed against various exposure indices using the Weibull model defined in Appendix A. Ranking and selection of exposure indices were based on minimizing the residual sum of squares (RSS) for each case. A relative RSS, calculated as the ratio of the RSS for each index to the minimum RSS for each case, was averaged across cases to calculate an average score. The selection criteria for an appropriate index were: (1) minimize the average scores, (2) minimize the range of relative RSSs across cases, and (3) minimize the percent variation (i.e., range/mean score * 100).

The results of the analysis indicated that while no single exposure index was best for all species, the best overall fits were obtained using indices that:

1. cumulated hourly concentrations over time,
2. placed greater weight on concentrations of 0.06 ppm or higher, either by continuous sigmoid weights or by discrete 0/1 weights, and
3. weighted hourly concentrations according to the phenological stage of plant development.

This is consistent with the evidence that exposure duration is important in determining plant response, and that repeated peaks are more important than mean concentrations in eliciting a response.

The results indicated that while the generalized phenologically weighted cumulative impact (GPWCI) indices best related plant response to O_3 exposure, there were other near optimal indices. These indices included the sigmoid-weighted integrated index (SIGMOID centered at 0.062 ppm), and the cumulative censored indices that integrated concentrations of 0.06 (or 0.07) ppm or higher (SUM06 and SUM07). Of all these best indices, only the SUM06 and SUM07 indices were suitable for regulatory purposes. The GPWCI and SIGMOID indices were deemed to be too complex for use as ambient air quality standards.

Mathematically, the SUM06 and SUM07 indices weight peak concentrations by discrete 0/1 step-functions that are first approximations to the more complex, continuous, sigmoid function centered at 0.062 ppm (used by the SIGMOID index).

EXPOSURE LEVELS ASSOCIATED WITH SPECIFIED CROP YIELD LOSSES

The Weibull exposure-response models, using HDM2 (the current standard), M7, SUM06, and SUM07 were used to predict relative yield losses and estimate the exposure levels of varying yield losses. The information is needed to develop strategies for limiting crop yield losses upon respecification of the secondary standard for ozone.

Ozone-induced crop yield loss was highly variable across species/cultivars, sites, and years. Crop yield losses for 16 NCLAN studies are illustrated in Figure 1.1. The horizontal lines of the boxplots are the 25th, 50th, and 75th percentiles of estimated yield losses across the 16 NCLAN cases. The magnitudes of plant responses are depicted by the 50th and 85th percentiles of predicted relative yield losses (PRYLs), which represent crop losses for typical and more sensitive crops, respectively.

Estimated crop losses for the 16 NCLAN studies ranged from 3.9% to 34.8% (interquartile range, i.e., 25th to 75th percentile), with a median crop loss of 18.9% in response to O_3

2

exposures having an HMD2 value of 0.12 ppm. Lowering the concentration to 0.10 ppm would result in an estimated median crop loss of 12.8%. To protect against a 10% (median) yield loss, the HDM2 standard would need to be set to 0.09 ppm.

Alternative standards for limiting crop reductions to 10% (median) could be achieved using the cumulative censored indices. Restricting exposures with SUM06 to less than 21 ppm-h (or SUM07 less than 18 ppm-h) would protect against a 10% (median) reduction.

Based on the median response of the 16 NCLAN cases (including O_3 tolerant crops), limiting exposure levels of M7 to 0.06 ppm would result in typical crop losses of 10% or less. However, neither the long-term averages nor the single peak indices are biologically relevant. The HDM2 and M7 indices do not adequately describe the temporal variations in exposure and do not relate well to plant response. A major shortcoming of noncumulative indices is the exclusion of exposure duration, an important factor in determining plant response.

AMBIENT AIR QUALITY ANALYSIS

Key issues of concern for the secondary ambient O_3 air quality standard are (1) the aggregation periods of exposure (relevant in defining the integrated indices, and less so for HDM2 and M7), and (2) the relationships among the current secondary NAAQS (HDM2), the 7-h seasonal mean (M7), and two integrated indices (SUM06 and SUM07).

Air quality analysis of 83 non-urban site-years of ambient O_3 data indicates that an aggregation period of 5 months from May to September (153 days) should be used to capture 90% of the hourly concentrations of 0.06 ppm or higher. This period includes the EPA-defined O_3 season for most sites, represents the period when the second highest daily maximum 1-h concentration generally occurs, and coincides with the agricultural season for the majority of crops grown in the United States.

An appropriate air quality indicator should correlate well with short-term peak, multiple peak, and long-term averages (U.S. EPA, 1987). Correlations of 0.75 or greater represent a strong association between indices. Fair to strong associations exist between the two cumulative censored indices (SUM06 and SUM07) and the peak and mean indices (HDM2 and M7). The integrated indices, SUM06 and SUM07, are strongly related to M7 (r=0.87 for both), and less related to HDM2 (r=0.67 and 0.71, respectively). The relationship between SUM07 and HDM2 (r=0.71) falls just below the defined level for a strong association.

Both SUM06 and SUM07 correlate well with a long-term index (M7) and a short-term index (HDM2) and relate well to biological response. These results suggest SUM06 and SUM07 have potential for setting a standard that protects against adverse effects from repeated peak and long-term exposures.

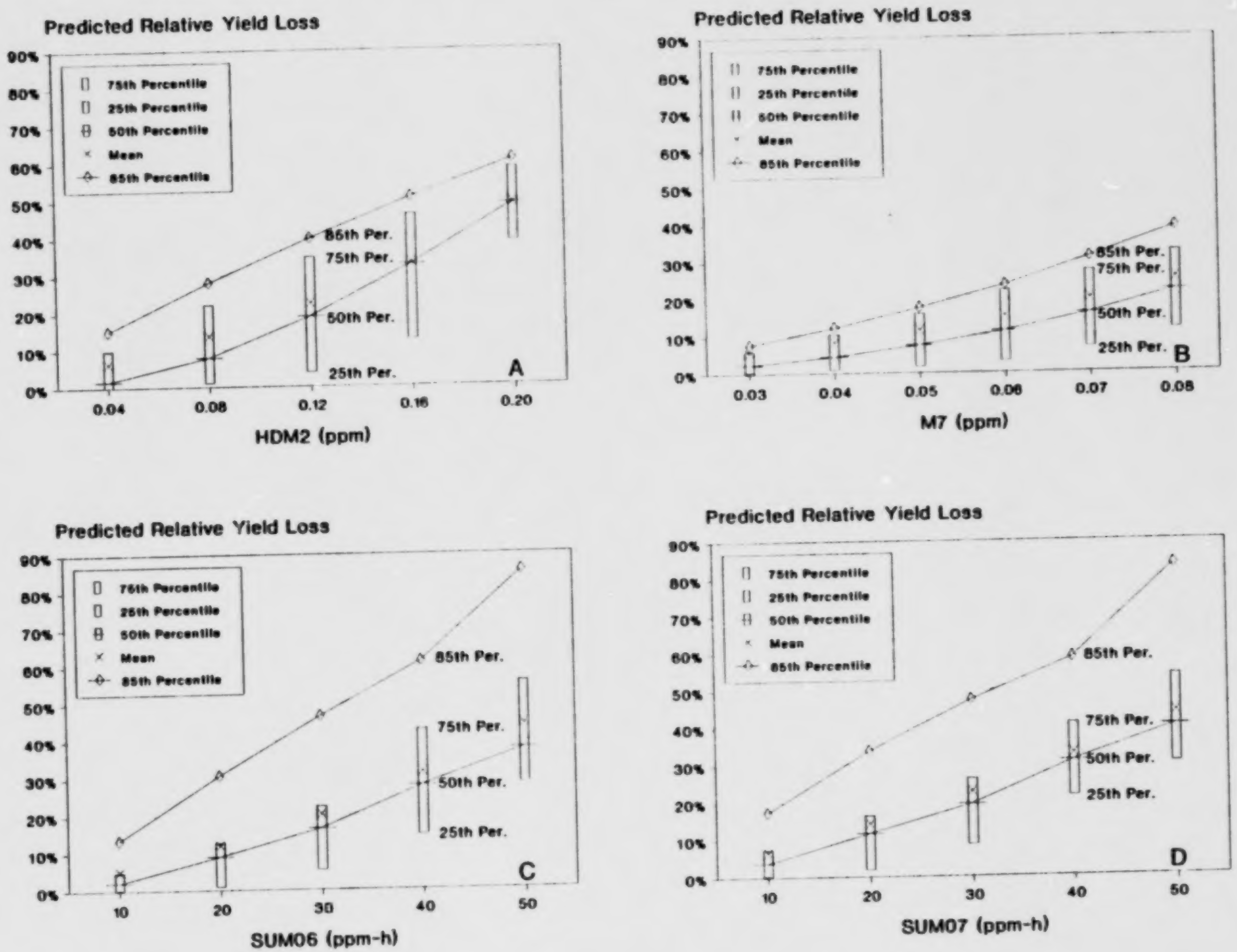


Figure 1.1. Boxplots of percentiles and mean predicted relative yield losses, characterizing the range of crop losses across the 16 NCLAN studies using the Weibull model with (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07 (same as Figure 3.6 in Chapter 3).

5

CHAPTER 2

COMPARISON OF OZONE EXPOSURE INDICES IN EXPOSURE-RESPONSE MODELING

2.1. INTRODUCTION

The goal of air pollutant research on crops is to determine the relationship between pollutant exposure and plant response. The quantifying function for this process is frequently termed "dose-response"; however, in this discussion of exposure indices it will be termed "exposure-response", a more generic concept. An understanding of exposure response requires three types of information: (1) a measure of plant response, (2) an appropriate index to describe the pollutant exposure, and (3) a mathematical function that relates plant response to pollutant exposure. Defining the appropriate exposure index that best relates plant response to exposure requires determining the underlying biological basis for the response and a defining method for characterizing the temporal variations in pollutant occurrence.

Plants react differentially to temporal variations in pollutant concentrations (e.g., Musselman et al., 1983; Hogsett et al., 1985, 1988; U.S. EPA, 1986); consequently, it is necessary to select exposure indices that specifically account for the temporal variation in exposure. If plant response is not modified by the episodic nature (temporal variation) of the exposures, then it is sufficient to characterize the exposure with an index, such as a mean, that does not reflect the concentration variation over time.

In evaluating exposure indices that characterize plant response, the ultimate goal is to develop exposure indices that incorporate all contributing features and account for all (or most) of the variation in the exposure response. A second, more practically oriented goal is the development of exposure indices useful in the ambient air quality standard setting process. An index for an ambient air quality standard should be easy to develop and applicable to a wide range of species and environmental/exposure conditions. The attainment of this objective, however, may represent a compromise in the features included in the formulation of the best exposure index.

The magnitude of plant response to O_3 is modified by several factors, including (1) climatic and edaphic factors, (2) biological factors, such as the stage of plant development and genetic potential, and (3) the temporal variation in O_3 exposure. These components and their influence on plant response were recently reviewed (U.S. EPA, 1986; Hogsett et al., 1988);

consequently, this analysis focuses on the evaluation of exposure indices that best relate O_3 exposure to plant response.

2.2. EVALUATION OF EXPOSURE INDICES

There is no consensus concerning the most appropriate exposure index for depicting plant response to O_3 exposure (Heagle and Heck, 1980; Hogsett et al., 1988). Different exposure indices have been used, but their adequacy for characterizing long-term exposures (over a season) is doubtful, since they do not account for exposure dynamics (i.e., temporal variation). In addition, the indices do not account for other known influences, such as the effects of exposure at specific and perhaps critical periods in plant development. Most indices, such as a seasonal mean, do not cumulate the impact of the exposure over the growing period.

An appropriate exposure index should ultimately be derived from a consideration of the underlying biological basis for the response and a procedure for quantifying the temporal variations in O_3 occurrence. An evaluation of various exposure indices is essentially a comparison of the influence of different scalings of the exposure variable on the exposure response function. The comparison is based on a single set of biological responses (e.g., plant yield) but they are related to different mathematical characterizations of O_3 exposure (scales), such as a mean or a cumulation of hourly concentrations. The different indices or scalings do not change the magnitude of the measured plant response, but they do govern how these responses are positioned along the exposure axis.

A pair of examples illustrates the effect of using different indices for characterizing the temporal variations in O_3 exposure (Figure 2.1). In the first example (Figures 2.1A-B), alfalfa plants were exposed to episodic O_3 exposures (Hogsett et al., 1985) and harvested (when the plants reached one-tenth bloom) three times during the season. A comparison was made between two exposure indices: (1) the 7-h (0900 to 1600) seasonal mean O_3 concentration (Heck et al., 1984a) and (2) the generalized phenologically weighted cumulative impact (GPWCI) that cumulated the exposure and emphasized the peak concentrations (Lee et al., 1988b). The fit of the Weibull exposure-response function to the identical plant response data was clearly better when the GPWCI (Figure 2.1B) was used to array the observed plant responses along the exposure axis than when the 7-h seasonal mean was used (Figure 2.1A). With the GPWCI, the trend for reduced plant growth with increasing exposure is plainly seen; however, this pattern is not evident when the mean is used. A second example (Figures 2.1C-D) is provided by a comparison of the maximum yield loss observed in a range of crop species. The evaluation of exposure indices compared the 7-h seasonal mean and the total O_3 exposure (sum of concentrations throughout the study). In this example, as in the previous one (Figures 2.1A-B), there is a clear trend showing an increase in maximum yield loss with increasing exposure (Figure 2.1D); however, when the 7-h seasonal mean was used this trend was not evident (Figure 2.1C). The principal difference between the two indices is the inclusion of exposure duration (number of days) in the total exposure index. Several of the studies had essentially the same 7-h seasonal mean, but the exposure durations differed by approximately twofold, which caused the lack of a clear trend in the data. In both examples, the mean was not an appropriate index (scaling factor) for describing the response of the plants to long-term (e.g., growing season) exposures to episodically occurring concentrations of O_3.

The mathematical mean of the hourly O_3 concentrations, over various time periods, is frequently used to characterize pollutant exposures. The 7-h (M7) and 12-h (M12) seasonal means have gained credence for relating plant yield to O_3 exposure because of their use in the National Crop Loss Assessment (NCLAN) program (Heck _et_ _al_., 1982, 1984a; Cure _et_ _al_., 1986). The specific 7-h period (0900-1600) selected was thought to correspond to the period of greatest plant sensitivity and the highest O_3 levels (Heck _et_ _al_., 1984a). The seasonal mean concentration is a summation of all hourly concentrations (for the selected time period) divided by the number of observations. Cure _et_ _al_. (1986) reported that seasonal mean characterizations of O_3 exposure were much less sensitive than the 1-h maximum to yearly variations in O_3 patterns. An infinite number of hourly distributions (from those containing many peaks to those containing none) can yield the same 7-h seasonal mean. If a mean is selected as an exposure index for characterizing exposures, certain assumptions that are implied and must be considered in a study's interpretation:

- A mean tends to minimize the contributions of peak concentrations to the response, implying that peak events do not need to be given special consideration.
- All concentrations within the selected averaging time are equally effective in eliciting a response.
- Reduced crop yield results from the accumulation of daily O_3 impacts over the growing season.
- Exposure duration is not specifically included; the mean cannot distinguish between exposures to the same concentration with different durations (e.g., 10 or 100 days).

- The distributions of hourly O_3 concentrations over the averaging time are unimodal and are not highly skewed. However, ambient O_3 concentration distributions are frequently skewed toward the higher concentrations.

Despite the shortcomings of a long-term average concentration (Tingey, 1984; Lefohn and Tingey, 1985), the 7-h seasonal mean has been proposed for use as a secondary ambient air quality standard for O_3 (Larsen *et al*., 1989). The authors calculated the correlations between 14 ambient ozone air quality indicators for 320 agricultural/remote SAROAD site-years and an estimated soybean crop reduction. The effective mean (Larsen and Heck, 1984) and M7 indices were best correlated with crop yield losses estimated from a lognormal model. No biological data were used in the analysis. Rather, the lognormal model that expressed crop reduction as a function of the effective mean concentration (Larsen and Heck, 1984) was used to generate crop loss estimates for the 320 site-years of ambient data. Since there was no biological variation in the data, correlations between the exposure indices and the estimated crop reductions were, in fact, measures of association between the (transformed) effective mean and the other indices. Consequently, no evidence that the mean indices were better correlated with plant response than other indices can be inferred from the analysis of Larsen *et al* (1989).

Previous analyses of plant-response data from NCLAN and other research programs found that exposure indices that best related to plant response were those that (1) cumulated hourly concentrations over time and (2) gave greater weights to higher concentrations (Lee *et al*., 1987, 1988b; Lefohn *et al*., 1988). These results support the importance of peak concentrations and the cumulative impact of ozone in relating plant response to exposure.

This report evaluates previously studied indices and some new multi-component indices that take into account peak concentrations, the cumulative impact of multiple peaks, and the biological response at varying stages of phenological development. Particular attention has been given to comparing sigmoid (Lefohn *et al*., 1988) and allometric (Lee *et al*., 1987) weights, and gamma and exponential phenological weights.

2.3. METHODS

Ideally, the selection of the most appropriate exposure index would be derived from scientific principles and be validated by experimental studies specifically designed for that purpose. However, this approach would delay the selection of exposure indices for years until the necessary experiments had been designed and conducted. An alternative approach is to conduct a retrospective analysis of existing plant-response data. This approach permits the reanalysis of existing data in order to evaluate or develop a range of exposure indices and derive principles of exposure response. Nevertheless, the indices developed by this method will ultimately need to be validated experimentally. Retrospective analysis faces a major limitation, however, in that the specific studies were not designed for developing exposure indices or testing exposure hypotheses; consequently, the usability of the data is limited (Rawlings et al., 1988).

2.3.1. EXPOSURE-RESPONSE DATA

Plant yield and hourly ozone monitoring data from 12 field experiments were obtained from the NCLAN Data Library in North Carolina:

1. Argonne, Illinois

    A. Soybean (Kress and Miller, 1983)
    B. Winter wheat (Kress et al., 1985)
    C. Corn (Kress and Miller, 1985a)
    D. Sorghum (Kress and Miller, 1985b)

2. Raleigh, North Carolina

    A. Soybean (Heagle et al., 1983, 1986)

3. Shafter, California

    A. Cotton (Temple et al., 1985)

4. Ithaca, New York

    A. Winter wheat (Kohut et al., 1987).

The 12 NCLAN studies represented five species with differing cultivars over four sites and multiple years for a total of 17 individual cases. Crops were grown according to standard agricultural practices and exposed to a range of ozone concentrations in open-top chambers, according to NCLAN protocols (Heck et al., 1983). Originally, two exposure studies (1983 wheat at Ithaca and 1981 soybean at Raleigh) involved a mixture of two pollutants, O_3 and SO_2, but were analyzed only for ozone using the ambient SO_2 chambers when an SO_2 effect was found.

2.3.2. EXPOSURE INDICES

Exposure indices, using hourly mean ozone concentrations, were calculated for the 7-h period (0900 to 1600 h) and classified into six distinct categories, according to their general characteristics (Tingey *et al*., 1988). Of the general measures previously used to describe ozone exposures, the numerical indices used in this analysis were grouped in the following manner:

1. *One Event*: included the current secondary NAAQS, that is the second highest daily maximum 1-h concentration (HDM2), the maximum of 7-h (P7) and 1-h (P1) maximum daily averages, and the 90th (PER90), 95th (PER95), and 99th (PER99) percentiles of hourly distribution.

2. *Mean*: included the seasonal mean of 7-h daily means (M7), the seasonal mean of 1-h daily peaks (M1), and the effective mean (EFFMEAN) (Larsen and Heck, 1984).

3. *Cumulative*: included the seasonal sum of hourly concentrations, i.e., total exposure (TOTDOSE).

4. *Concentration Weighting*: two subclasses of weighting indices were defined:

    a. *Sum Concentration* (Sum Conc): included the seasonal sum of hourly concentrations at or above 0.06 ppm (SUM06), 0.07 ppm (SUM07), 0.08 ppm (SUM08; see Figure 2.2), or 0.10 ppm (SUM10); the seasonal censored sum of hourly concentrations at or above 0.08 ppm (AOT08; see Figure 2.2), or 0.10 ppm (AOT10); the total impact (TIMPACT) (Larsen *et al*., 1983); the ALLOMETRIC in which the hourly concentration was raised to a power and summed; the SIGMOID in which the hourly concentration was variably weighted using a sigmoid function and summed (see Appendix A).

    b. *Number of episodes* (Num Episodes): included the total hours with concentrations at or above 0.08 ppm (HRS08), or 0.10 ppm (HRS10); the number of episodes (an episode was defined as an event with hourly ozone concentrations above a threshold level) of threshold 0.08 ppm (NUMEP08), or 0.10 ppm (NUMEP10); and the average episode length with threshold of 0.08 ppm (AVGEP08), or 0.10 ppm (AVGEP10).

5. *Multicomponent*: indices that incorporated several characteristics of exposure, including the phenologically weighted cumulative impact indices (PWCI) (Lee *et al*., 1987).

6. Respite Time: included the average number of days between episodes of 0.08 ppm (DAYBET08) or 0.10 ppm (DAYBET10).

In addition, a generalization of the PWCI index originally developed by Lee et al. (1987) was used to generate a set of peak-weighted cumulative indices; the notation O(i) denotes the hourly concentration, W[O(i)] denotes a continuous peak-weighting function, PW(t) denotes a weight that serves as a surrogate for plant sensitivity at day t, and F[O(i), O(i-1)] denotes a weight for intra-episodic concentrations. Then

$$\mathrm{GPWCI} = \Sigma(W[O(i)] * PW(t) * F[O(i),O(i-1)] * O(i)).$$

The general shapes of the weighting functions, illustrated in Figures 2.3A-D, were dictated by biological considerations and formulated into mathematical expressions (see Appendix A).

For W[O(i)], eight sigmoid (SIG1 to SIG8) and seven allometric (AL1 to AL7) functions were used to continuously weight the hourly concentrations. The eight sigmoid functions covered a wide range of possible weight functions and were centered between 0.048 and 0.104 ppm (see Figure 2.3A). At the lower extreme (SIG6 - centered at 0.048 ppm), hourly concentrations larger than 0.048 ppm were given weights of 0.5 or greater (maximum weight equals 1); at the upper extreme (SIG2 - centered at 0.104 ppm), hourly concentrations below 0.08 ppm were given weights of 0.15 or smaller (Figure 2.4).

The seven allometric functions were originally used by Lee et al. (1987) to emphasize peak concentrations (Figure 2.3B). In comparing those 15 weighting functions, the allometric function AL6 was much less sensitive to hourly concentrations below 0.08 ppm (Figure 2.4) than the sigmoid function SIG6. The allometric function AL1 de-emphasized large concentrations, whereas AL2 represented the uniform weighting (Figure 2.3B) used in calculating a mean or the total exposure.

Crop sensitivity to ozone varies, depending on the biological stage of plant development. The magnitude of ozone-induced plant response depends, in part, on the timing of peak concentrations relative to plant phenology. Since the relationships between phenology and sensitivity to O_3 are not well understood, various phenological weight functions PW(·) were considered (Figure 2.3C). The constant function (PW01), two exponential functions (PW02, PW03), and 16 gamma functions (PW04 to PW19), were used as surrogates of plant sensitivity curves. Function PW01 treated all stages of plant development equally, with no times of increased plant sensitivity. The exponential functions weighted ozone concentrations as though the plants were most sensitive at crop maturity, whereas the gamma functions placed the time of maximum sensitivity prior to crop maturity. The first eight gamma functions, PW04 to PW11, had zero weight at final harvest, with maximum weights occurring 5 to 40 days before crop maturity (Figure 2.3C). The 40 days before final harvest represented half or more of the exposure days for most crops. The remaining eight gamma functions, PW12 to PW19, were generated by shifting the first eight gamma functions to the right by 10 days (Figure 2.5). The shifted gamma functions had non-zero weights at maturity, with maximum weights occurring 0 to 30 days before maturity. All phenological weighting functions are scaled to a maximum weight of 1.

Three intra-episodic schemes (F1, F2, F3) for adjusting the peak concentrations and their effect within an episode were compared (Figure 2.3D). The episodic schemes were selected to represent various changes in pollutant uptake that a plant might employ. In Figure 2.3D, hourly concentrations were unaffected by F1, but adjacent concentrations (after the initial peak concentration) at or above 0.08 and 0.10 ppm were decreased by a factor of 0.8 for F2 and F3, respectively. Non-uniform weights decreased the effect of later peak concentrations within an episode, possibly reflecting the plant's biological mechanisms during an episode.

At the initial peak concentration, the plant's defense mechanism, via stomatal closure, decreases the effect of further acute concentrations by limiting pollutant uptake. The weighting schemes, F2 and F3, place less weight on later peaks to reduce the effective "dose" within an episode. The allometric weighted GPWCI indices were calculated by using only the uniform weight, F1.

2.3.3. WEIBULL EXPOSURE-RESPONSE MODEL

Regression models are extremely useful for quantifying the relationship between air pollution stress and biomass production, but they do not provide an explanation of the biochemical and physiological processes in O_3-stressed plants. To assess the economic impact of ozone on plant growth, the Box-Tidwell (Kopp et al., 1984) and Weibull (Rawlings and Cure, 1985) models have been used to relate plant response to exposure. A comparison of the two models (Lee et al., 1988a) favors the Weibull model, because (1) the Weibull model parameter estimates and associated exposure levels are more stable and more precise than the Box-Tidwell estimates and (2) the Weibull function has a horizontal asymptote of zero, which extrapolates to zero growth in the limit. However, when these models were used to estimate ozone effects on crop yield or to evaluate exposure indices, similar results were obtained using either the Box-Tidwell or the Weibull model (Lee et al., 1988a). For comparison of exposure indices, we reported the relationships between plant response and ozone exposure using the Weibull model.

The Weibull curve is a fairly simple monotonically decreasing function with maximum value at zero. Assuming an additive error structure with a covariate (Z), the Weibull model relates plant response (Y) to an ozone exposure index (X) by

$$Y = (A + D*Z) * \exp[-(X/B)^{C}] + e. \tag{1}$$

where e is the model error. The Weibull model, fit to log transformed data and excluding the covariate, is

$$\ln(Y) = A' + B'*X^{C'} + e. \tag{2}$$

The model parameters in (2) are reparametrizations of the model parameters in (1) with D=0, i.e., $A'=\ln(A)$, $C'=C$, and $B'= -(1/B)^{C}$.

2.3.4. REGRESSION APPROACH TO COMPARE EXPOSURE INDICES

In total, crop yield data from 12 NCLAN studies, with multiple varieties for a total of 17 cases, were regressed against 25 general and 589 GPWCI indices using the Weibull model. Regressions of harvest yields for each chamber on the 614 ozone exposure indices were performed using nonlinear least squares in SAS (1985).

The selection criterion for determining the "best" index was based on the minimum residual sum of squares (RSS) for each of the 17 cases. Ranking and selection of exposure indices were performed separately for each case. For each individual case, the relative performance of an index was measured as the ratio of its RSS to the minimum RSS. Indices were ranked in ascending order; that is, the "best" index, for a specific case, had a relative RSS of 1. Kendall's coefficient of concordance (C_T) and Pearson's correlation coefficient, computed on Savage scores, were used to measure the agreement in the rankings across the 17 cases (Iman and Conover, 1987). These statistics are sensitive to agreements in the top rankings and have a maximum value of 1 when there is perfect agreement in rankings. The asymptotic chi-squared approximation for Kendall's statistic was used to test the independence among the various sets of rankings.

For overall comparison among species and cultivars, the exposure indices were evaluated according to three criteria: (1) the average score calculated as the mean relative RSS's averaged across the 17 cases, (2) the range of relative RSS's, and (3) the percent variation (i.e., range/mean score x 100). For each case, the exposure indices were ranked according to the average score and classified into fairly homogeneous groups of 100. (The last group had 114 indices.) The nonparametric Wilcoxon signed ranks test was used to perform pairwise comparisons of relative RSS's among two indices across cases (Conover, 1971). Individual comparisons were not made among all possible pairs to avoid the multiple comparison problems of controlling for error rate. Pairwise comparisons between the top-ranked index and a select number of indices were calculated. When significance was found, it was assumed that lower-ranked indices were significantly different from the top-ranked index.

2.4. RESULTS

Based on the criterion of minimum RSS, no single exposure index performed best for all 17 cases (representing the different species and studies), but there was a positive agreement in the top rankings among the 17 cases (C_T=0.035). Restricting attention to the 25 general indices and the 285 GPWCI indices with F1 (the uniform weight of intra-episodic concentrations), the agreement among the top rankings was much stronger with a Kendall's coefficient of 0.101. Kendall's coefficient is significantly different from 0 (P-value<0.01). These statistical tests indicate that although no index always had the minimum RSS, there was a group of indices that were always at or near the minimum RSS. Given the positive agreement in top rankings among the cases, the remaining analysis focused on identifying key characteristics in determining plant response, based on the average score and the variation in relative fits across species, sites, and years.

Of the 614 ozone exposure indices, the GPWCI index, PWCI485, with sigmoid weight SIG3 (centered at 0.062 ppm), phenological gamma weighting scheme PW10 (maximum weight at 30 days prior to maturity), and intra-episodic weighting function F3 (adjusted down adjacent concentrations of 0.10 ppm and higher), was overall best, with individual scores ranging from 1.02 to 1.36 with an average score of 1.12 (minimum) (Table 2.1A).

Based on average scores across the 17 cases, other GPWCI indices were near optimal, including the 100 top performing indices, having a typical average score of 1.14. Classification of the indices in sets of 100 resulted in mean average scores of 1.16, 1.18, 1.20, 1.22, and 1.82 for the second to sixth groups, respectively. The paired sample Wilcoxon signed ranks test showed no differences among the top-ranked index and the 300th ranked index (P-value >0.05). It was concluded that there were no significant differences among the 300 top performing indices with average scores between 1.12 and 1.19. However, there was a significant difference in relative RSS's between the top-ranked GPWCI index and the 350th ranked index (P-value<0.05); it was concluded that the top-ranked index performed significantly better than the 260 bottom rankings, including M7, which had average scores of 1.20 or greater.

In comparison with the best GPWCI, 21 of the 25 general indices, excepting SUM07, SUM06, SUM08, and AOT08, had average scores of 1.20 or higher, placing them in the bottom three groups

(Table 2.1B). Restricting attention to the 25 general indices and the top-ranked GPWCI index, Kendall's concordance coefficient C_T was 0.08 (P-value=0.15), indicating weak, positive agreement in top rankings among the 17 cases, but when the 1982 wheat study at Ithaca (which had convergence problems in the nonlinear fits to the various exposure indices) was excluded, Kendall's C_T improved to 0.11 (p-value = 0.01). The paired sample Wilcoxon signed ranks test showed no significant differences, at the 0.05 level, between the top-ranked index and other indices that gave greater weight to peak concentrations and cumulated the exposures (SUM07, SUM06, SIGMOID, SUM08, and AOT08). The weighted cumulative indices, SUM07, SUM06, and SIGMOID, had similar average scores of 1.15, 1.15, and 1.17, respectively, and displayed similar percent variations to the top-ranked GPWCI index across all species/cultivars.

Although the SUM08 and AOT08 indices were not significantly different from the top performing index, they displayed greater percent variation than the top performing GPWCI index. Based on the three criteria, SUM07, SUM06, and SIGMOID were nearly equal in performance to the top-ranked GPWCI index, and were significantly better than M7.

Of the 589 GPWCI indices, the sigmoid-weighted GPWCIs accounted for 186 of the 200 top performing indices, whereas the GPWCIs with allometric weights were not among the 100 top performing indices (Table 2.2). Sigmoid weights centered between 0.048 and 0.073 ppm gave better fits to plant response than did sigmoid functions centered at 0.083 ppm or higher (Table 2.2).

Compared to the sigmoid-weighted GPWCI indices, the GPWCIs with allometric weights did not relate well to plant growth (Table 2.2). A slight increase in performance for increasing values of the allometric parameter was noted.

Table 2.3 shows the distribution patterns of the GPWCI indices by phenological weighting function. The top performing GPWCI indices used gamma functions with maximum weights occurring 20 to 40 days before harvest. These indices accounted for 88 of the 100 top performing exposure indices (Table 2.3). When the unshifted (zero weight at crop maturity) and shifted gamma functions (right shift of 10 days such that maximum weight occurs 10 days near harvest) were compared, the unshifted gamma functions were found to perform better than their shifted versions in all cases. Those GPWCI indices with phenological weighting schemes having the greatest weight near the time of final harvest (e.g., the exponential functions) were among the poorer performing indices. On the other hand, the GPWCIs with uniform phenological weighting schemes had above average performance but accounted for only 10 of the top 100 indices.

Finally, the sigmoid-weighted GPWCI indices were classified by intra-episodic weighting schemes. For the sigmoid-weighted GPWCI indices, slightly better fits were obtained using the intra-episodic weighting scheme, F3 (i.e., a weighting that reduced weight given to the latter portion of an O_3 event, simulating reduced O_3 uptake), than were obtained using F1 or F2 (Table 2.4); 65 of the 200 top indices use F3, versus 62 and 59 using F2 and F1, respectively.

The overall best exposure index was the GPWCI index with sigmoid weight SIG3 and phenological gamma function PW10 (and intra-episodic function F3), but other combinations were near optimal. The choice of the intra-episodic scheme was less important than the selection of sigmoid and phenological weights. In particular, near optimal scores (averaged across the 3 intra-episodic schemes) of 1.13 to 1.14 were obtained for those indices with sigmoid and phenological weights of (SIG3,PW08), (SIG4,PW08), (SIG6,PW08), (SIG7,PW08), (SIG8,PW08), (SIG3,PW11), (SIG6,PW11), (SIG7,PW11), (SIG8,PW11), (SIG3,PW16), and (SIG7,PW16); see Table 2.5. These top performing GPWCI indices had sigmoid weights with inflection points between 0.056 and 0.073 ppm, and had phenological weighting schemes with a maximum occurring 20 to 30 days prior to crop maturity. Performance rankings of the GPWCI indices for combinations of sigmoid and phenological weights were little affected by the choice of the intra-episodic weights.

2.5. DISCUSSION

The analysis indicates that those indices that emphasize peak concentrations and cumulate concentrations over time perform better than those that average concentrations. Similar conclusions were reached by Lefohn _et_ _al_. (1988) and Lee _et_ _al_. (1987, 1988b), who used NCLAN data and cumulative indices with sigmoid and allometric weights. Of the two methods for emphasizing peaks, the current analysis found the sigmoid-weighted GPWCI indices substantially outperformed the allometric-weighted indices. The GPWCI indices with sigmoid weights also performed better than the censored cumulative exposure indices, which ignored hourly concentrations below thresholds of 0.08 and 0.10 ppm. However, near optimal performances were obtained using the weighted cumulative indices, SUM07, SUM06 and SIGMOID (centered at 0.062 ppm). These were similar to the top-ranked GPWCI indices in emphasizing concentrations of 0.07 (0.06) ppm or greater and accumulating concentrations over time.

Of the six classes of exposure indices evaluated, best fits were obtained using the indices classified as multicomponent (GPWCI) or sum concentration weighting (SUM07, SUM06, SIGMOID) indices. Based on our criteria, the noncumulative indices classified as one event, mean, number of episodes, or respite time were significantly inferior to the weighted cumulative indices. These results suggest that ambient ozone levels below 0.08 ppm are important in triggering plant response and should be included in an exposure index. Indices that placed little (or equal) weight on lower concentrations, such as the allometric-weighted indices with allometric parameter values of 2 or greater (Figure 2.4), and the seasonal mean had performances much below the top-ranked indices.

The threshold indices are simple and easy to implement, from a regulatory perspective, although there is little biological evidence to suggest a threshold exists or is constant among phenological stages or species. Considering that SAROAD monitoring sites in the United States experience ambient ozone between 0.05 and 0.06 ppm about 30% of the time (U. S. Environmental Protection Agency, 1982), appreciable yield losses may be undetected using the threshold indices, i.e., that have a high threshold, e.g., 0.08 ppm, but will be identified with the sigmoid-weighted exposure indices or threshold indices that have a low threshold (e.g., 0.06 ppm).

2.6. CONCLUSION

These analyses support the conclusions from previous studies that demonstrated the importance of peak concentrations in determining plant response. Although no single index was deemed best (in all cases) for relating O_3 exposure to plant response, the top performing exposure indices were those that (1) cumulated the hourly O_3 concentrations over time, (2) emphasized concentrations of 0.06 ppm and higher either by continuous sigmoid weights or by discrete (0 or 1) weights of the threshold indices, and (3) gave greater weight to exposures occurring 20 to 40 days before harvest. When assessing the impact of O_3 on plant growth, these findings illustrate the importance of exposure duration, the importance of repeated peaks, and the time of increased plant sensitivity.

TABLE 2.1

A. Performance criteria for 25 general ozone exposure and 3 other types of indices (GPWCI, SIGMOID, ALLOMETRIC).

Exposure index	Mean score	Minimum score	Maximum score	Range	% Variation	Index type
GPWCI (best)	1.12	1.02	1.36	0.34	30	Multicomponent
SUM07	1.15	1.03	1.31	0.28	25	Sum Concentration
SUM06	1.15	1.03	1.37	0.34	30	Sum Concentration
SIGMOID	1.15	1.03	1.33	0.30	26	Sum Concentration
SUM08	1.17	1.04	1.56	0.52	45	Sum Concentration
AOT08	1.18	1.02	1.81	0.79	67	Sum Concentration
TOTDOSE	1.20	1.02	1.58	0.56	47	Cumulative
M7	1.20	1.02	1.58	0.56	47	Mean
TIMPACT	1.20	1.01	1.66	0.65	54	Sum Concentration
EFFMEAN	1.20	1.01	1.67	0.66	55	Mean
M1	1.21	1.01	1.82	0.81	67	Mean
PER90	1.22	1.00	1.81	0.81	66	One Event
PER95	1.23	1.01	1.94	0.93	76	One Event
ALLOMETRIC	1.22	1.01	2.00	0.99	80	Sum Concentration
PER99	1.26	1.00	2.21	1.21	96	One Event
AOT10	1.26	1.00	2.85	1.85	147	Sum Concentration
SUM10	1.28	1.04	2.98	1.94	152	Sum Concentration
P7	1.31	1.00	2.38	1.38	105	One Event
HDM2	1.33	1.10	1.92	0.82	61	One Event
HRS10	1.44	1.00	4.65	3.65	254	Num Episode
HRS08	1.46	1.04	5.04	4.00	274	Num Episode
AVGEP08	1.47	1.01	4.90	3.89	265	Num Episode
P1	1.61	1.00	4.88	3.88	241	One Event
AVGEP10	1.72	1.00	4.06	3.06	178	Num Episode
NUMEP10	1.80	1.00	5.43	4.43	245	Num Episode
NUMEP08	2.09	1.00	5.64	4.64	222	Num Episode
DAYBET10	24.43	2.39	99.27	96.88	397	Respite Time
DAYBET08	37.94	2.39	253.77	251.38	663	Respite Time

Note: Paired sample Wilcoxon signed ranks tests found no significant differences (at the 0.05 level) between the top performing index and the following indices: SUM07, SUM06, SIGMOID (inflection point at 0.062 ppm), SUM08, and AOT08.

B. Classification of 25 general ozone exposure indices ranked by average score in groups of 100 (Weibull model).

Group	Frequency	Ozone exposure indices						
1-100	2	SUM07	SUM06					
101-200	1	SUM08						
201-300	1	AOT08						
301-400	4	TOTDOSE	M7	TIMPACT	EFFMEAN			
401-500	3	M1	PER90	PER95				
>501	14	PER99	AOT10	SUM10	P7	HDM2	HRS10	HRS08
		AVGEP08	P1	AVGEP10	NUMEP10	NUMEP08	DAYBET08	DAYBET10
Total	25							

TABLE 2.2

Classification of the 589 ozone exposure indices ranked by average score in groups of 100, and classified according to the sigmoid and allometric functions (Weibull model).

	Sigmoid weighting functions								
Group	SIG1	SIG2	SIG3	SIG4	SIG5	SIG6	SIG7	SIG8	Subtotal
1-100			22 (39%)	15 (26%)		19 (33%)	21 (37%)	21 (37%)	98
101-200	18 (32%)		8 (14%)	15 (26%)	19 (33%)	10 (18%)	9 (16%)	9 (16%)	88
201-300	11 (19%)	5 (9%)	9 (16%)	9 (16%)	11 (19%)	5 (9%)	8 (14%)	9 (16%)	67
301-400	10 (17%)	25 (44%)	3 (5%)	3 (5%)	9 (16%)	5 (9%)	4 (7%)	3 (5%)	62
401-500	10 (17%)	13 (23%)	9 (16%)	9 (16%)	12 (21%)	8 (14%)	7 (12%)	9 (16%)	77
>500	8 (14%)	14 (25%)	6 (11%)	6 (11%)	6 (11%)	10 (18%)	8 (14%)	6 (11%)	64
Total	57	57	57	57	57	57	57	57	456
c	70	70	100	100	100	130	130	130	
M	500	1500	500	1500	4000	500	1500	4000	
Inflection point	0.089	0.104	0.062	0.073	0.083	0.048	0.056	0.064	

Allometric weighting schemes

Group	AL1	AL2	AL3	AL4	AL5	AL6	AL7	Subtotal
1-100								0
101-200	2 (11%)	2 (11%)	2 (11%)	2 (11%)	1 (5%)	1 (5%)	1 (5%)	11
201-300	2 (11%)	3 (16%)	4 (21%)	5 (26%)	6 (32%)	6 (32%)	6 (32%)	32
301-400	3 (16%)	5 (26%)	4 (21%)	5 (26%)	5 (26%)	6 (32%)	6 (32%)	34
401-500	3 (16%)	2 (11%)	3 (16%)	2 (11%)	3 (16%)	3 (16%)	4 (21%)	20
>500	9 (47%)	7 (37%)	6 (32%)	5 (26%)	4 (21%)	3 (16%)	2 (11%)	36
Total	19	19	19	19	19	19	19	133
Allometric power	0.5	1.0	1.5	2.0	2.5	3.0	3.5	

TABLE 2.3

Classification of the 589 ozone exposure indices ranked by average score in groups of 100, and classified according to the phenological weighting gamma functions (Weibull model).

	Phenological weighting functions											
Group	PW01	PW02	PW03	PW04	PW05	PW06	PW07	PW08	PW09	PW10	PW11	
1-100	10					14		15	3	15	14	
101-200	8			15		7		10	13	8	11	
201-300	3			6		7	14	4	9	5	4	
301-400	10			9		3	11	2	5	3	2	
401-500		25		1	22		5		1			
>500		6	31		9		1					
Total	31	31	31	31	31	31	31	31	31	31	31	
a	1.0	1.0	1.0	1.5	1.5	2.0	2.0	2.5	2.5	3.0	3.0	
b	0.00	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10	
End	LAST	LAST	LAST	LAST	LAST	LAST	LAST	LAST	LAST	LAST	LAST	
Time of maximum[a]		70	70	60	65	50	60	40	55	30	50	

	Phenological weighting functions								
Group	PW12	PW13	PW14	PW15	PW16	PW17	PW18	PW19	Subtotal
1-100					12		15		98
101-200			14		7		6		99
201-300	11		7		7		8	14	99
301-400	12		9		5	12	2	11	96
401-500	6		1	15		16		5	97
>500	2	31		16		3		1	100
Total	31	31	31	31	31	31	31	31	589
a	1.5	1.5	2.0	2.0	2.5	2.5	3.0	3.0	
b	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10	
End	SHIFT	SHIFT	SHIFT	SHIFT	SHIFT	SHIFT	SHIFT	SHIFT	
Time of maximum	70	75	60	70	50	65	40	60	

[a] Time of maximum weight is calculated assuming a 70-day exposure period.

TABLE 2.4

Classification of the 589 generalized phenologically weighted cumulative impact indices ranked by average score in groups of 100, and classified according to the intra-episodic weights (Weibull model).

	Intra-episodic weights					
Group	F1 with allometric function	F1 with sigmoid function	F2 with sigmoid function	F3 with sigmoid function	Subtotal	Average score
1-100		30 (20%)	30 (20%)	38 (25%)	98	1.14
101-200	11 (8%)	29 (19%)	32 (21%)	27 (18%)	99	1.16
201-300	32 (24%)	24 (16%)	19 (13%)	24 (16%)	99	1.18
301-400	34 (26%)	22 (14%)	24 (16%)	16 (11%)	96	1.20
401-500	20 (15%)	26 (17%)	24 (16%)	27 (18%)	97	1.22
>500	36 (27%)	21 (14%)	23 (15%)	20 (13%)	100	1.82
Total	133	152	152	152	589	
Average score	1.21	1.19	1.19	1.18		1.30

TABLE 2.5

Average scores for various combinations of the sigmoid and phenological gamma functions for calculating the generalized phenologically weighted cumulative impact indices (Weibull model).

Phenological weighting functions	Sigmoid weighting functions								Average
	SIG1	SIG2	SIG3	SIG4	SIG5	SIG6	SIG7	SIG8	
PW01	1.17	1.19	1.15	1.15	1.17	1.16	1.15	1.15	1.17
PW02	1.22	1.23	1.21	1.21	1.21	1.22	1.21	1.21	1.22
PW08	1.16	1.18	1.13	1.14	1.16	1.14	1.13	1.13	1.15
PW09	1.18	1.19	1.15	1.16	1.18	1.16	1.15	1.15	1.17
PW11	1.17	1.19	1.14	1.15	1.17	1.14	1.14	1.14	1.16
PW16	1.17	1.19	1.14	1.15	1.16	1.15	1.14	1.16	1.16
Average	1.20	1.21	1.18	1.18	1.19	1.19	1.18	1.18	

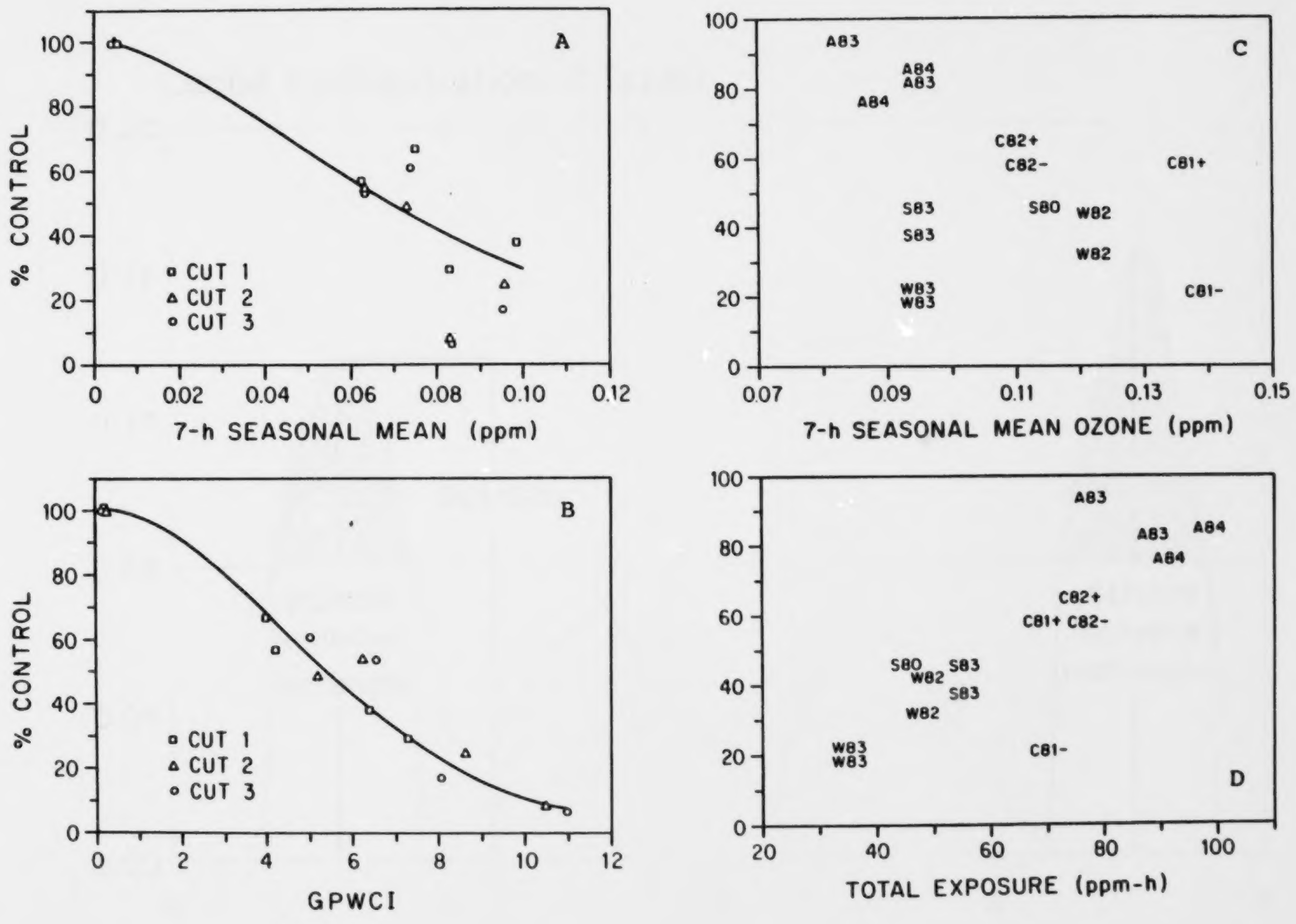


Figure 2.1. Two comparisons of different exposure indices using the same biological response data. % Control = Treatment yield/Control yield x 100 where the Control is defined as the yield in charcoal-filtered air. The data in Figure 2.1A and B are derived from Hogsett et al. (1988) and the data in Figure 2.1C and D are from Lee et al. (1987). In Figure 2.1C and D, the symbols indicate the year of the study and the species used; A = alfalfa, C= cotton, S = soybean, W = wheat.

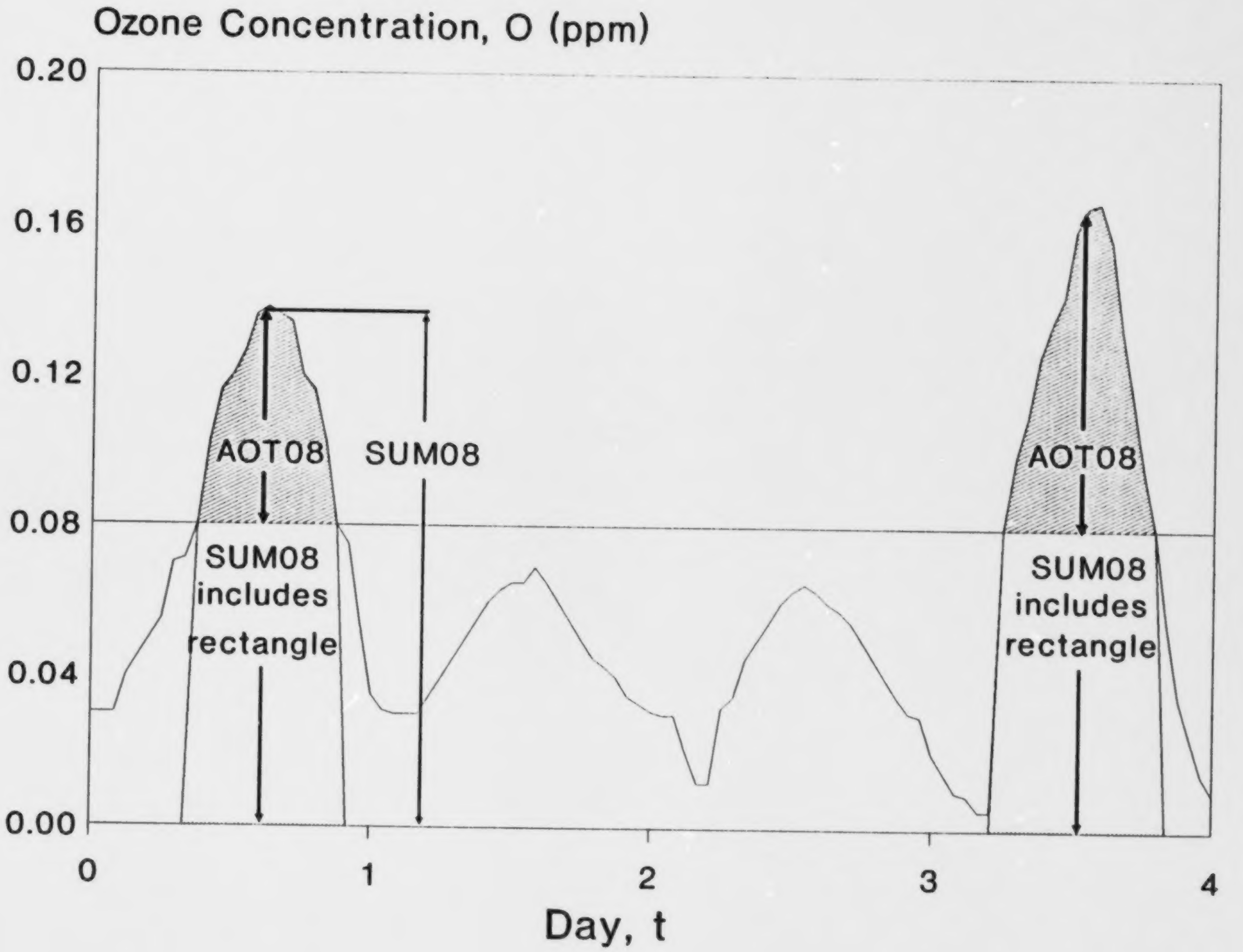


Figure 2.2. Description of the methods used to calculate the areas under hourly concentration curve for the cumulative indices, AOT08 and SUM08.

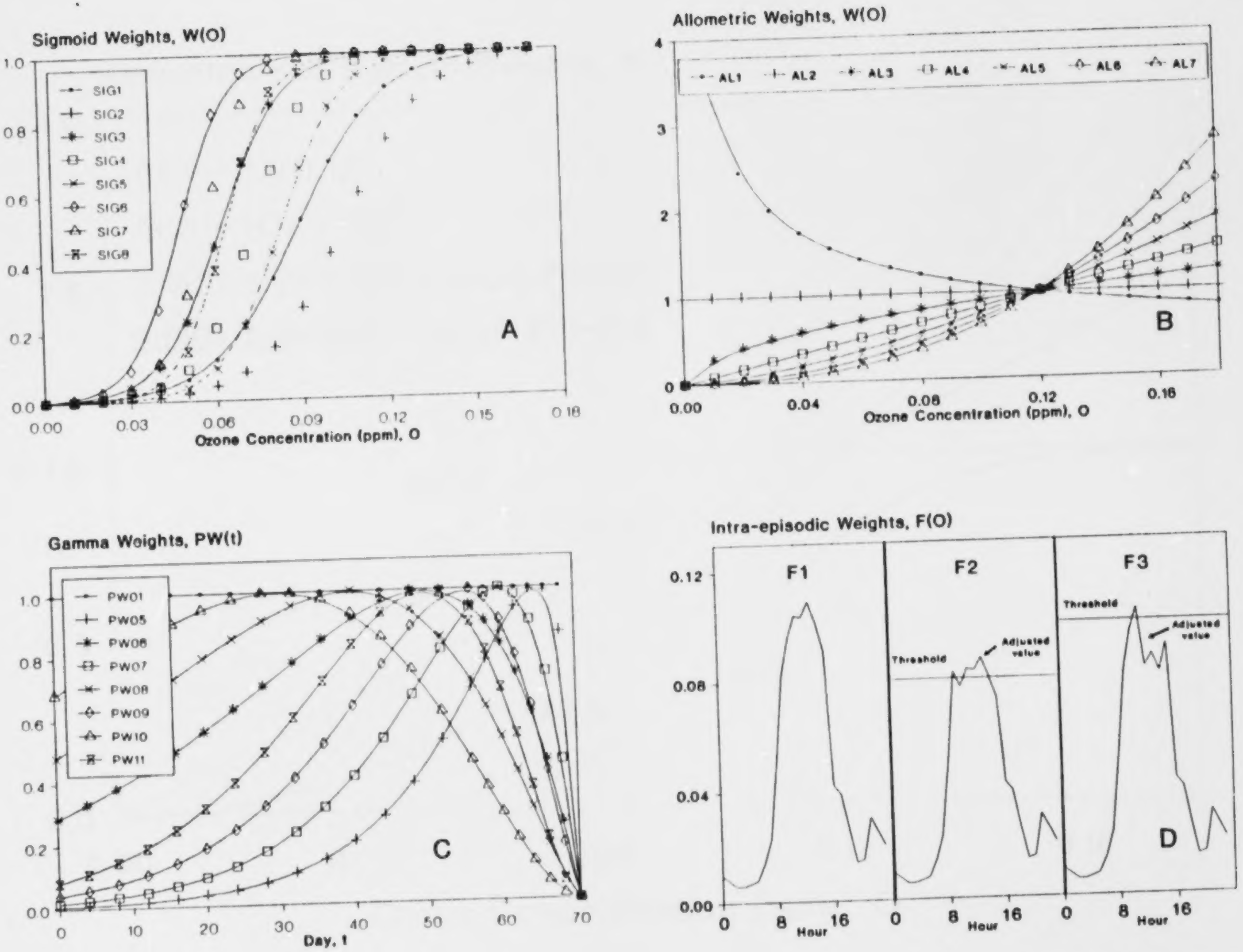


Figure 2.3. Weighting functions used to calculate the generalized phenologically-weighted cumulative impact index: sigmoid (A) and allometric (B) functions for weighting the hourly ozone concentrations; (C) gamma functions mimicking plant sensitivity over time (assumes a 70-day exposure); and (D) three intra-episodic weights mimicking pollutant uptake during an episode.

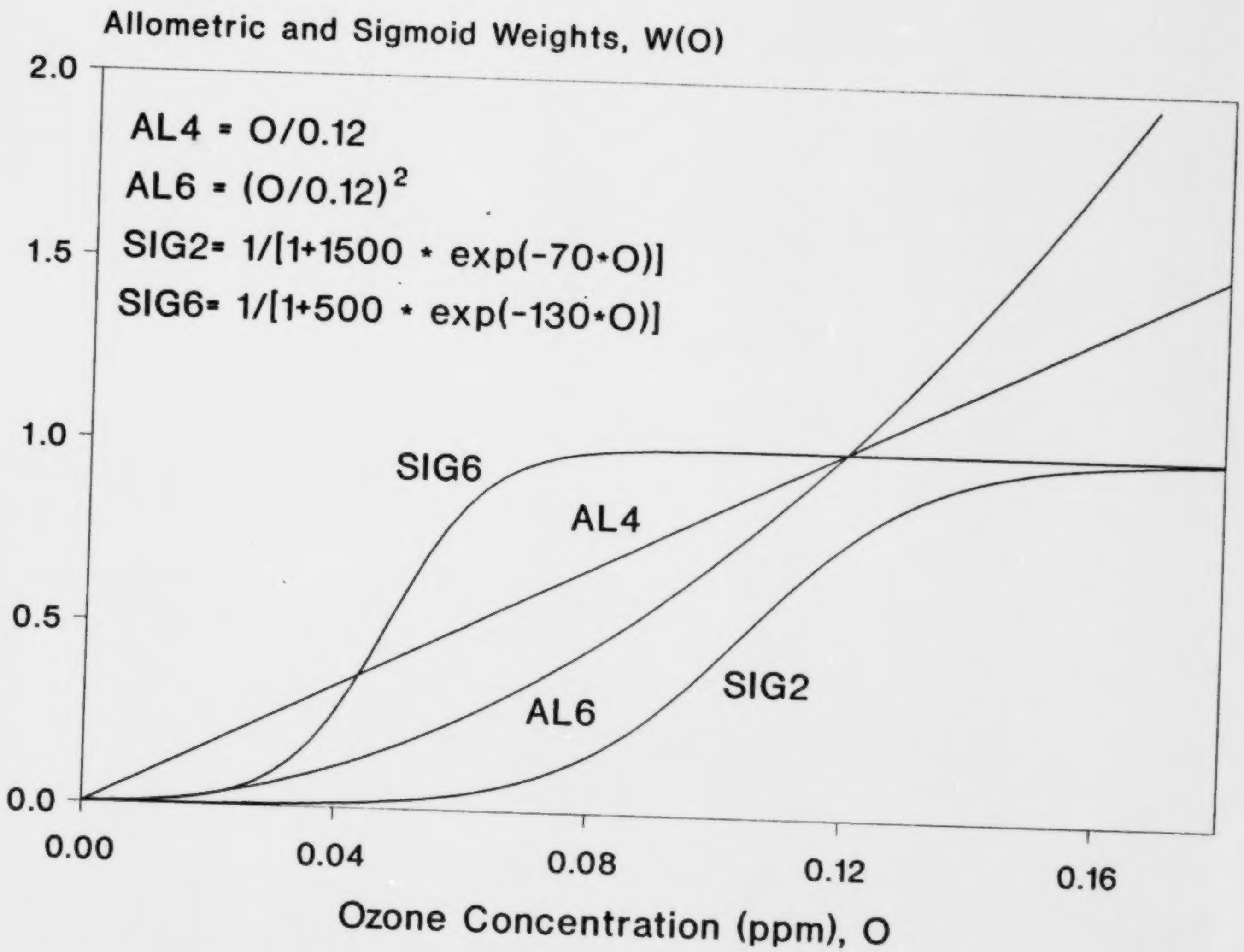


Figure 2.4. Allometric and sigmoid functions used to weight the hourly ozone concentrations. For comparison purposes, the allometric functions were scaled to have weights of 1 at ozone concentration of 0.12 ppm.

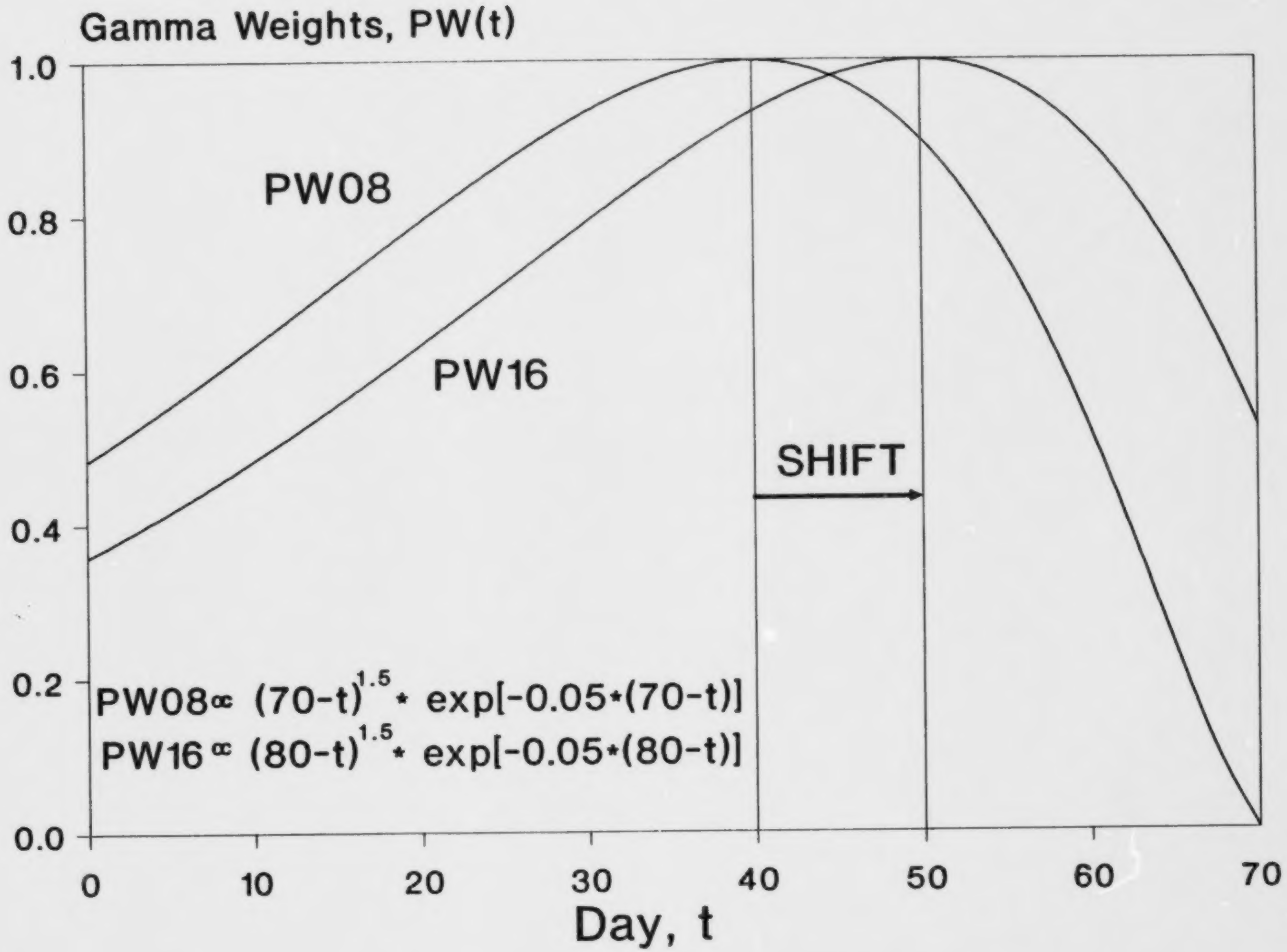


Figure 2.5. Examples of mathematical functions used to simulate changes in plant sensitivity with time. The shifted and unshifted phenologically-weighted gamma functions were used to weight the hourly ozone concentrations over time (assuming a 70-day exposure period).

CHAPTER 3

ESTIMATION OF EXPOSURE LEVELS ASSOCIATED WITH VARIOUS CROP YIELD LOSSES

3.1. INTRODUCTION

To support regulatory efforts to develop an appropriate ambient air quality standard for O_3, it is necessary to quantify the magnitude of crop loss expected to occur at various exposure levels. The exposure-response models developed in Chapter 2 were used to predict relative yield losses (PRYLs) and estimate the exposure levels associated with specified crop yield losses for various crops. Because the evaluation of various exposure indices (Chapter 2) found that no single index was best for relating O_3 exposure to plant yield for all crops/species, sites, and years, a range of exposure indices (HDM2 - current NAAQS, M7, SUM06, SUM07) was used.

The evaluation of exposure indices (Chapter 2) found that the GPWCI index best depicted the O_3 effects on plant yield; however, other indices, SIGMOID, SUM06, and SUM07, were near optimal. Although an exposure index is highly correlated with crop loss, it does not follow that it has a form suitable for specifying an ambient air quality standard. For example, the GPWCIs, the best indices (Chapter 2), are crop-dependent, including a phenological weighting term that changes with crop development, and hourly O_3 concentrations are weighted using a sigmoid function. Although biologically relevant, the GPWCI indices are too complex for developing an ambient air quality standard. Similarly, the Office of Air Quality Planning and Standards determined that the SIGMOID index was inappropriate for developing an ambient air quality standard. They felt that the computational steps were too complex for monitoring agencies to implement in the field.

The cumulative censored (i.e., "threshold") indices, SUM06 and SUM07, are adequate descriptors of exposure; they relate well to biological response and are simple and easy to implement from a regulatory perspective. Mathematically, the SUM06 and SUM07 indices are similar to the SIGMOID index that uses a sigmoid function centered at 0.062 ppm. The sigmoid function (centered at 0.062 ppm) is an S-shaped curve that increases monotonically from 0 to 1 and equals 0.5 at 0.062 ppm. The discrete 0-1 step function, with a discontinuity at 0.06 or 0.07 ppm, is a first approximation of the continuous S-shaped sigmoid curve. That is, the area under the sigmoid curve is approximately equal to the area under the 0-1 step function. The threshold indices are simplifications of the SIGMOID index.

The objective of this chapter is to quantify the magnitude of crop yield losses associated with varying exposure levels, as quantified with the SUM06, SUM07, M7, or the HDM2. Based on the exposure-response models derived in Chapter 2, exposure levels of the four exposure indices associated with yield losses of 5%, 10%, 15%, 20%, 25% and 30% were calculated.

3.2. METHODS

To estimate percent crop loss at various levels of O_3 exposure, plant response data for wheat, soybean, corn, sorghum, and cotton were regressed against various exposure indices (HDM2, M7, SUM06, and SUM07), using Weibull exposure-response models. The plant response data included 16 separate cases involving five species, and/or multiple cultivars or years for the species. Crop yield data for the two cotton and two corn studies were log transformed to stabilize experimental variability.

In the NCLAN field experiments, which provided the plant response data, the plants were exposed to a range of O_3 exposures varying from charcoal-filtered (CF) to levels above ambient. The range of experimental O_3 concentrations and the duration of exposures for the NCLAN studies are shown in Table 3-1. Exposure duration ranged from 36 to 111 days with an average of 75 days. The O_3 exposures were started at varying periods following seedling emergence and continued to maturity.

The form of the Weibull model relating plant growth (Y) to an exposure index (X) and a possible covariate (Z) is

$$Y = (A+D*Z)*\exp[-(X/B)^{C}] + e, \qquad (1)$$

where e is the error term and A, B, C, and D are unknown model parameters. For log transformed data, the Weibull model can be expressed in a linearized form, namely,

$$\ln(Y) = A'+D'*Z+B'*X^{C'} + e, \qquad (2)$$

where A', B', C', and D' are unknown model parameters. Ignoring the error terms and covariate, the models in (1) and (2) are mathematically equivalent, i.e., $A'=\ln(A)$, $B'=-(1/B)^{C'}$, and $C'=C$.

For the Weibull model fit to untransformed data, the predicted relative yield loss (PRYL) associated with an exposure level of (X) is

$$\mathrm{PRYL}(X) = 1 - \exp[-(X/B)^{C}]. \qquad (3)$$

Conversely, for a fixed yield loss of s% (i.e., setting the left-hand side of equation (3) to s%), the exposure level (X) is

$$X(s) = B*[-\ln(1-s/100)]^{(1/C)}. \qquad (4)$$

For log transformed data, the parameter B' is reported in Table 3.2. We calculate $B=(-1/B')^{(1/C)}$ for use in (3) and (4).

For example, the Weibull model fit using M7 for the 1982 cotton study has estimates of -112.2 and 2.1451 for parameters B' and C. We calculate $B = (1/112.2)^{(1/2.1451)} = 0.1107$. A 10% exposure level for M7 is X(10) = 0.039 ppm from (4).

A background O_3 concentration of 0.025 ppm (expressed as the M7) was used in the NCLAN studies (e.g., Heck *et al*., 1983, 1984b) to calculate relative yield losses. In this analysis, crop-yield losses were estimated relative to zero O_3 for all indices. For all four exposure indices, over the O_3 concentration range from 0 to the O_3 concentration in charcoal-filtered air, the Weibull exposure-response curves are essentially flat, such that relative yield losses reported using either 0 or an assumed background (i.e., 0.025 ppm for M7) will not differ appreciably (e.g., Weibull model for the 1983 soybean study with Amsoy at Argonne, IL; see Figure 3.1). Our analyses indicate yield differences are less than 2% between zero and charcoal-filtered levels.

Because the 16 NCLAN studies represent species/cultivars both O_3 sensitive and tolerant, magnitudes of crop yield losses for specified values of the four exposure indices are described using the 25th, 50th, 75th, and 85th percentiles and the mean PRYLs. The median (i.e., 50th percentile) is less influenced by extreme losses than the mean and represents the typical crop yield loss. That is, half the crop losses lie below the median and half lie above it.

3.3. RESULTS

The magnitude of the O_3-induced yield losses varied among sites, years, and species/cultivars. The 16 NCLAN studies used in this analysis represent both O_3 sensitive and tolerant species/cultivars, and multiple years and experimental sites. The data set included six wheat studies representing three cultivars (Abe, Arthur, Vona), two sites (Argonne, IL and Ithaca, NY) and two years of data at each site. There were five soybean studies representing three cultivars (Amsoy, Corsoy, and Davis), two sites (Argonne, IL and Raleigh, NC), and two years at each site. Corn, cotton, and sorghum studies were included and represented a single site (Argonne, IL or Shafter, CA) for one or two years.

Initially, relative yield losses at four values per index were calculated using equation (3) and the Weibull parameters in Table 3.2. Figures 3.2-3.5 illustrate the range of crop loss that can occur in various species/cultivars under field conditions.

In these figures, ABE 82 and ARTHUR 82 refer to the 1982 wheat studies at Argonne (A82WH); ABE 83 and ARTHUR 83 refer to the 1983 wheat studies at Argonne (A83WH); CORSOY 80, AMSOY 83, and CORSOY 83 refer to the 1980 and 1983 soybean studies at Argonne (A80SO, A83CV), respectively; DAVIS 81 and 82 refer to the 1981 and 1982 soybean studies at Raleigh (R81SO, R82SO), respectively; PAG 81 and PIONEER 81 refer to the 1981 corn study at Argonne (A81MA); SJ-2 81 and 82 refer to the 1981 and 1982 cotton studies at Shafter (C81CO, C82CO); VONA 82 and 83 refer to the 1982 and 1983 wheat studies at Ithaca (I82WH, I83WH), respectively; and DEKALB 82 refers to the 1982 sorghum study at Argonne.

Plant response to O_3, as described by the Weibull model, is a monotonically decreasing function that predicts decreased crop yield with increasing exposure levels (e.g., Figure 3.1) for all four exposure indices. To illustrate the range in PRYL at various levels of the exposure indices, the PRYL function was solved for specific levels of each index (Figures 3.2, 3.3, 3.4, 3.5). For each level of an exposure index, a wide range in the crop yield loss estimates, among species/cultivars and site-years, was evident. Sorghum and corn were less sensitive to O_3 exposure. On the other hand, the wheat cultivar, Vona, was extremely sensitive to O_3 for both years. In Vona wheat, there was greater variation in PRYLs when the noncumulative indices HDM2 and M7 (Figures 3.2, 3.3) were used than when yield loss was estimated using the SUM06 and SUM07 (Figures 3.4, 3.5) exposure indices. The lower year-to-year variability in PRYLs when the cumulative indices (SUM06 and SUM07) were used was largely explained by differences in exposure durations (71 and 36 days for the 1982 and 1983 studies, respectively). The calculated value of the cumulative exposure indices (SUM06 and SUM07) was influenced by exposure duration while the HDM2 and M7 were insensitive to differences in exposure duration. A similar pattern in variation among exposure indices was observed in soybean (Davis 81 and Davis 82).

The other cultivars of wheat (Abe, Arthur) were less sensitive to O_3 than Vona. At SUM07=20 ppm-h, multiple years of data from Arthur showed similar yield losses but Abe was less O_3 sensitive in 1983 than in 1982 (Figure 3.5C). Soybean cultivars, Corsoy and Davis, showed yield losses as large as or larger than the wheat cultivars, Abe and Arthur. However, multiple years of Corsoy show varying O_3 sensitivities in 1980 and 1983 using M7, SUM06, or SUM07 (Figure 3.3-3.5) but not using HDM2 (Figure 3.2).

Two years of cotton display similar crop yield losses using M7, SUM06, or SUM07 but not using HDM2. Yield losses for cotton plants were comparable to yield losses experienced by soybean (Davis).

To summarize crop yield losses into a single plot, the PRYLs for each of the 16 studies were presented as boxplots (Figure 3.6). In the boxplots, the horizontal lines denote the 25th, 50th, and 75th percentiles of PRYLs for the 16 studies. For example, at the 50th percentile, half the 16 studies showed smaller yield losses (at a given value of the exposure index) and the other half displayed larger losses. Similarly, at the 75th percentile, three-quarters of the studies showed smaller yield losses and one-quarter of the studies showed greater losses. The 50th and 85th percentiles of PRYLs at various levels of the four exposure indices were joined to allow linear interpolation of exposure levels. These summary values appear in Table 3.3. The upper levels of the four exposure indices were chosen based on (1) the highest experimental O_3 exposures given in Table 3.1 and (2) the highest ambient O_3 exposures for agricultural/remote sites (described more fully in Chapter 4).

Figure 3.6A indicates that the relative yield losses associated with HDM2=0.12 ppm vary considerably (range of the 25th to 75th percentiles) with an estimated median PRYL of 18.9%, for the 16 NCLAN cases. Dropping the level to 0.10 ppm for HDM2 would result in an estimated median yield reduction of 12.8%, a 32% improvement.

Figure 3.6B shows the relative yield losses associated with M7 values between 0.02 and 0.08 ppm. There is less variation (range of the 25th to 75th percentiles) in PRYLs using M7 than using HDM2. The median PRYLs of 10% and 20% occur at exposure levels of approximately 0.06 and 0.08 ppm for the M7. The highest M7 value in non-urban areas in the United States is less than 0.08 ppm (Chapter 4).

Figures 3.6C-D show the PRYLs associated with SUM06 and SUM07 values ranging from 10 to 50 ppm-h. There is greater variation in PRYLs using SUM06 than using SUM07. At levels of 50 ppm-h or lower, the SUM07 gives higher losses than SUM06.

The key features of Figures 3.6A-D were summarized by calculating the exposure levels, for each of the exposure indices, associated with 5%, 10%, 15%, 20%, 25% and 30% PRYLs. Linear interpolation of the 50th, 75th, and 85th percentiles and the mean PRYLs is used to calculate the exposure levels (Table 3.4). Typical losses are characterized by the 50th percentile (median) PRYL. To protect against a 10% yield loss, the HDM2 index must be set at 0.089 ppm, the M7 index at 0.057 ppm, the SUM06 at 21.3 ppm-h, or the SUM07 at 17.9 ppm-h.

3.4. DISCUSSION

In our calculations, the median PRYLs were used to estimate the associated O_3 exposure levels. This approach would protect against typical crop yield losses but not necessarily against the maximum losses that could occur. Alternative strategies for determining exposure levels associated with prespecified yield losses that are more protective could use higher (e.g., the 75th or 85th) percentiles of PRYLs rather than the median PRYLs. A strategy between the median and 85th percentile might be more appropriate, considering the variation in responses across the 16 NCLAN studies. A probabilistic approach could be used to determine the appropriate percentile that is relevant to setting a standard.

An important consideration in developing a standard is the selection of study crops and the relative weight of selected species and cultivars in terms of economic, agricultural, or biological significance. To assess the impact of ozone on agricultural crops, the NCLAN program selected (1) crops of major economic importance, (2) the most commonly used cultivars in a region, and (3) cultivars of intermediate sensitivity if known (Heck et al., 1984a). The study crops in our analysis included the three major cash crops in the corn belt, corn, soybean, and wheat, and represented a substantial share of commercial crops. We attached equal importance to each species and cultivar to describe the magnitudes of biological response from ozone exposure and to develop strategies that protect the more sensitive crops against pollutant loss. Future analyses will include more species and cultivars to obtain a more complete representation of agriculture.

The impact of ozone on crops is measured in terms of biological response (crop yield). In turn, the yield loss estimates from the Weibull model can be input into a U.S. agricultural sector model to measure the economic effects of ozone on agriculture. Based on a U.S. agriculture sector model using NCLAN data, Adams et al. (1985) found the economic effects in terms of economic surplus to be less than the physical effects. For example, for a 25% reduction in ozone, they estimate a 2.9% increase in crop yields for all crops and a 1.4% increase in economic surplus. Consideration of the economic effects can aid in determining the tolerance level of crop loss for use in developing ozone control strategies.

The four exposure indices were calculated using hourly ozone concentrations for the 7-h period 0900-1600 h, and were unchanged (M7, HDM2) or increased marginally (SUM06, SUM07) when calculated using all hourly concentrations. The analysis of ambient ozone concentrations at the NCLAN sites found few hourly concentrations of 0.06 ppm or higher outside the 7-h period and similar PRYLs would be expected using the 24-h exposure indices. Few studies or experiments are available to determine an appropriate averaging time for an exposure index that considers the timing of peak concentrations relative to the diurnal fluctuations of plant activity and sensitivity to O_3. However, longer averaging times (than the 7-h period) are needed to capture all the elevated ambient O_3 concentrations at many sites in the United States.

The results of the current analysis differ significantly from those of Larsen et al. (1989). In the current analysis, actual air quality and plant response data were used in deriving the exposure-response functions. Separate exposure-response functions were derived for each exposure index. In contrast, Larsen et al., (1989) used ambient air monitoring data to derive a correlation between the M7 and other indices. They developed their exposure-response functions for only the "effective mean" and then used the correlation (based on ambient air monitoring data) to adjust the plant response data. Consequently, the lack of fit that they report is not the lack of fit based on a biological (plant) response but rather the variation associated with the lack of correlation between the "effective mean" and other exposure indices used to characterize ambient concentrations.

3.5. CONCLUSION

Generalizing our analysis indicates that median crop reductions of 18.9% can be expected in agricultural sites experiencing O_3 exposures having a HDM2 value of 0.12 ppm. A stricter standard of 0.10 ppm for the HDM2 index is needed to protect against typical crop losses of 12.8%. A level closer to 0.09 ppm is needed to protect against a 10% median crop loss.

Alternative standards for limiting crop reductions to 10% could be achieved using the cumulative censored indices. In terms of the SUM06 index, exposures should not exceed 21 ppm-h for a 10% yield loss standard. In terms of the SUM07 index, exposures should not exceed 18 ppm-h for a 10% yield loss standard. The aggregation period for SUM06 and SUM07 is defined in terms of the experimental exposure "seasons" used in the NCLAN studies. Relating the exposure levels associated with the cumulative censored indices to ambient air quality data require that the issue of aggregation period be addressed. In Chapter 4 and Appendix B, the temporal features of ambient air quality data for 83 non-urban sites in the United States are analyzed to interrelate the experimental and ambient air quality indicators.

35

TABLE 3.1.

Range of ozone treatments and exposure durations used in the 12 NCLAN studies.

ID	OZONE TRMT	M7 (ppm)	SUM06 (ppm-h)	SUM07 (ppm-h)	HDM2 (ppm)	# DAYS
A80SO	CF	0.022	0.3	0.0	0.065	56
	CA90	0.114	42.9	42.5	0.200	
A81MA	CF	0.014	0.2	0.0	0.059	83
	CA120	0.156	89.9	89.8	0.245	
A82SG	CF	0.016	0.0	0.0	0.053	85
	CA100	0.129	76.1	75.9	0.208	
A82WH	CF	0.023	0.1	0.0	0.055	56
	CA90	0.122	47.2	46.9	0.200	
A83CV	CF	0.023	0.2	0.0	0.062	83
	CA60	0.094	52.3	49.8	0.184	
A83WH	CF	0.018	0.1	0.0	0.056	53
	CA60	0.095	33.5	32.7	0.175	
C81CO	CF	0.018	0.0	0.0	0.033	72
	CA100	0.137	67.5	65.6	0.247	
C82CO	CF	0.012	0.0	0.0	0.036	98
	PA20	0.111	71.8	70.2	0.203	
I82WH	CF	0.021	0.0	0.0	0.053	71
	CA90	0.096	41.7	40.7	0.181	
I83WH	CF	0.027	0.9	0.4	0.075	36
	PA18	0.098	22.2	21.3	0.205	
R81SO	CF	0.024	0.9	0.4	0.087	111
	CA70	0.125	94.5	93.8	0.219	
R82SO	CF	0.019	0.3	0.1	0.065	90
	CA60	0.111	67.2	65.5	0.193	

Experimental treatments range from charcoal-filtered ozone exposures (CF) to incremental additions to ambient O_3 in non-filtered air either by constant addition of xx ppb (CAxx) or proportional addition by factor of x.x (PAxx).

TABLE 3.2.

Values of the parameters of the Weibull exposure-response functions for each of the 16 NCLAN study cases. The Weibull parameters are given for four exposure indices (HDM2, M7, SUM06, and SUM07).

MODEL PARAM.	A80SO CORSOY	A83CV AMSOY	A83CV CORSOY	R81SO DAVIS	R82SO DAVIS	A82WH ABE	A82WH ART	A83WH ABE	A83WH ARTHUR	I82WH VONA	I83WH VONA	A82SG DEKALB	A81MA[d] PAG-397	A81MA[d] PIONEER	CB1CO[d e]	CB2CO[d f]
Exposure Index = HDM2 (ppm)																
A or A'	3316.9	2233.0	2171.9	588.7	530.3	-1377.1	436.2	1223.2	3448.3	10790.9	5647.1	618.5	6.0891	6.0000	7.0900	5.2868
B or B'	0.2211	0.3087	0.2549	0.2207	0.2629	0.2328	0.2378	0.2040	0.2307	0.0770	0.2151	0.3883	-2577.1	-227.4	-4.18	-24.47
C	2.4874	1.3965	2.1319	1.0000	1.6360	2.5045	2.4880	12.51	5.9295	1.0000[g]	1.2702	2.9432	5.6530	3.8427	1.0406	2.0240
D or D'	-18.96[a]					70.76[b]	50.76[b]	29.23[c]	13.53[c]							
SSE	463893	1622436	862245	20365	27688	3226975	1465212	493990	1210508	5833627	16853484	18359	0.2616	0.1798	0.0668	0.4743
Exposure Index = M7 (ppm)																
A or A'	3259.3	1929.2	1905.6	451.3	490.2	-1317.1	507.0	1529.1	3708.6	8736.5	4678.2	618.3	6.0824	5.9866	6.9752	5.2823
B or B'	0.1328	0.1169	0.1057	0.1509	0.1232	0.1433	0.1473	0.1117	0.1411	0.0474	0.1140	0.2963	-2487.3	-187.1	-31.18	-112.2
C	2.0048	3.5935	6.2357	1.4358	2.3709	2.4593	2.3379	10.44	3.8588	1.0000[g]	2.0539	2.2181	4.2671	2.8169	1.7768	2.1451
D or D'	-19.28					69.10	49.25	27.33	11.87							
SSE	476521	1351671	708415	11220	17407	2739773	1238003	418489	1060876	4364767	16782283	19209	0.2864	0.1936	0.0428	0.4716
Exposure Index = SUM06 (ppm)																
A or A'	3094.6	1900.4	1888.0	419.5	480.5	-1187.2	478.5	1517.4	3577.1	4934.6	4459.8	611.1	6.0761	5.9674	6.9049	5.2271
B or B'	56.09	72.10	60.47	131.9	83.48	58.26	59.71	42.41	73.13	27.13	29.54	176.6	-5.0E-9	-4.1E-6	-4.4E-4	-0.0018
C	1.5130	2.4925	4.8496	1.2076	1.5129	2.1375	2.1465	7.28	2.0666	1.0000[g]	1.3294	2.2720	4.2208	2.7534	1.7887	1.4598
D or D'	-17.52					66.16	47.70	27.39	12.72							
SSE	432270	1383008	714692	8902	14966	2731056	1287694	417914	1094793	8816721	15695198	19062	0.2902	0.1829	0.0444	0.4466
Exposure Index = SUM07 (ppm)																
A or A'	3086.6	1919.2	1892.1	414.4	477.5	-1112.2	420.3	1509.0	3585.4	4622.7	4450.0	610.6	6.0713	5.9636	6.9023	5.2188
B or B'	60.26	80.65	60.562	144.9	88.23	61.97	63.60	44.64	86.09	26.86	30.26	200.4	-4.8E-9	-8.5E-6	-8.2E-4	-0.0042
C	1.2214	1.6445	3.5398	1.0000	1.2264	1.6923	1.7444	5.55	1.6674	1.0000[g]	1.1074	1.9796	4.2282	2.5863	1.6516	1.2611
D or D'	-17.38					65.23	48.20	27.44	12.57							
SSE	466367	1381523	709899	8462	14895	2849058	1373000	419978	1103800	12372235	16302576	18861	0.2919	0.1700	0.0443	0.4574

---

[a] For A80SO, a virus problem became manifest and the percent of unhealthy plants (x 100%) was included as a covariate (mean = 24.4; Kress and Miller, 1983).

[b] For A82WH, the number of heads/meter was included as a covariate (Kress et al., 1985). Covariate means are 96.3 (Abe) and 82.9 (Arthur).

[c] For A83WH, the number of heads/meter was included as a covariate (Kress et al., 1985). Covariate means are 166.7 (Abe) and 147.0 (Arthur).

[d] Log transformations of plant growth data for CB1CO, CB2CO, A81MA were fit using the Box-Tidwell model.

[e] For the 1981 cotton study (CB1CO), only well-watered plants were used in the nonlinear regressions.

[f] For the 1982 cotton study (CB2CO), no significant differences between drought-stressed and well-watered plants were found. A single Box-Tidwell model was fit to both drought-stressed and well-watered plants, ignoring the drought factor.

[g] For the 1982 wheat study at Ithaca (I82WH), the Weibull parameter, C, was set to one for convergence of the nonlinear routine (Heck et al., 1984b).

37

TABLE 3.3

Percentiles and mean predicted relative yield losses (PRYLs) associated with various levels of the four exposure indices, HDM2, M7, SUM06, and SUM07, for the 16 NCLAN cases. The distribution of PRYLs for the 16 NCLAN cases is characterized by the 25th, 50th, 75th, and 85th percentiles and the mean.

HDM2		Percentiles of PRYLs			
(ppm)	25th	50th	MEAN	75th	85th
0.04	0.1%	1.7%	6.3%	9.8%	15.0%
0.06	0.4%	4.2%	9.9%	15.9%	21.7%
0.08	1.1%	7.9%	13.7%	22.1%	28.0%
0.10	2.2%	12.8%	17.8%	28.8%	33.8%
0.12	3.9%	18.9%	22.3%	34.8%	39.7%
0.16	12.6%	32.6%	33.0%	46.0%	50.5%
0.20	38.6%	48.7%	48.6%	58.3%	60.4%

M7		Percentiles of PRYLs			
(ppm)	25th	50th	MEAN	75th	85th
0.02	0.1%	0.2%	3.4%	0.6%	1.4%
0.03	0.3%	2.3%	5.6%	5.9%	7.6%
0.04	0.9%	4.4%	8.3%	10.4%	12.3%
0.05	1.8%	7.5%	11.5%	16.0%	17.6%
0.06	3.1%	11.3%	15.3%	22.3%	23.5%
0.07	6.7%	15.9%	19.8%	27.2%	30.9%
0.08	11.5%	22.0%	25.1%	32.4%	38.7%

SUM06		Percentiles of PRYLs			
(ppm-h)	25th	50th	MEAN	75th	85th
10	0.2%	2.2%	5.1%	4.9%	13.4%
20	0.9%	9.0%	12.0%	12.6%	30.6%
30	5.4%	16.5%	20.2%	22.3%	46.5%
40	14.6%	27.9%	31.4%	42.8%	61.1%
50	28.2%	37.6%	43.8%	55.5%	85.3%

SUM07		Percentiles of PRYLs			
(ppm-h)	25th	50th	MEAN	75th	85th
10	0.3%	3.7%	6.7%	7.2%	17.2%
20	2.0%	11.7%	14.3%	16.4%	33.7%
30	8.6%	19.4%	22.7%	26.2%	47.4%
40	21.5%	30.9%	32.8%	40.9%	58.5%
50	30.2%	40.2%	43.8%	53.7%	83.4%

TABLE 3.4.

Exposure levels associated with predicted relative yield losses (PRYLs) of 5% to 30% for the four exposure indices, HDM2, M7, SUM06, and SUM07, for the 16 NCLAN studies. Separate calculations are performed using the 50th, 75th, and 85th percentiles, and the mean PRYLs for the 16 NCLAN studies.

A) Exposure levels with Mean PRYLs of 5 to 30%.

	Mean Predicted Relative Yield Loss					
	5%	10%	15%	20%	25%	30%
HDM2 (ppm)	0.032	0.061	0.087	0.110	0.131	0.149
M7 (ppm)	0.027	0.045	0.059	0.070	0.080	0.089
SUM06 (ppm-h)	9.7	17.1	23.7	29.8	34.3	38.8
SUM07 (ppm-h)	7.5	14.4	20.9	26.8	32.3	37.2

B) Exposure levels with 50th Percentile PRYLs of 5 to 30%.

	50th Percentile Predicted Relative Yield Loss					
	5%	10%	15%	20%	25%	30%
HDM2 (ppm)	0.065	0.089	0.107	0.123	0.137	0.152
M7 (ppm)	0.042	0.057	0.068	0.077	0.085	0.093
SUM06 (ppm-h)	14.1	21.3	28.0	33.1	37.5	42.2
SUM07 (ppm-h)	11.6	17.9	24.3	30.5	34.9	39.2

C) Exposure levels with 75th Percentile PRYLs of 5 to 30%.

	75th Percentile Predicted Relative Yield Loss					
	5%	10%	15%	20%	25%	30%
HDM2 (ppm)	0.023	0.041	0.057	0.073	0.089	0.104
M7 (ppm)	0.028	0.039	0.048	0.056	0.066	0.075
SUM06 (ppm-h)	10.2	16.6	22.4	27.6	31.3	33.7
SUM07 (ppm-h)	6.9	13.0	18.5	23.7	28.8	32.6

D) Exposure levels with 85th Percentile PRYLs of 5 to 30%.

	85th Percentile Predicted Relative Yield Loss					
	5%	10%	15%	20%	25%	30%
HDM2 (ppm)	<0.020	0.026	0.040	0.055	0.070	0.087
M7 (ppm)	0.026	0.035	0.045	0.054	0.062	0.069
SUM06 (ppm-h)	3.7	7.5	10.9	13.8	16.7	19.6
SUM07 (ppm-h)	2.9	5.8	8.7	11.7	14.7	17.8

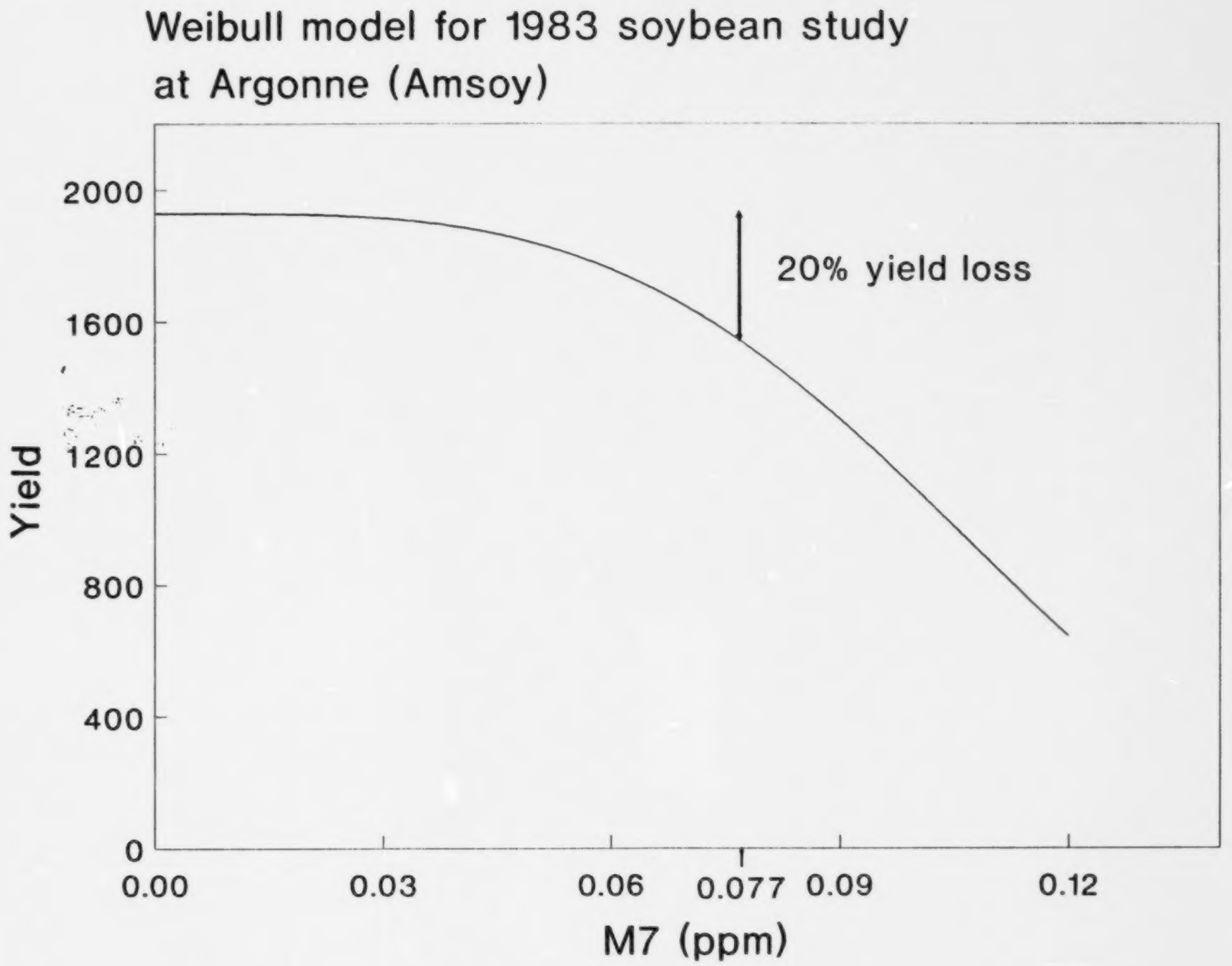


Figure 3.1. Illustration of the exposure level associated with a 20% relative yield loss using the exposure-response model for 1983 soybean (Amsoy) study at Argonne, IL with the 7-h seasonal mean (M7).

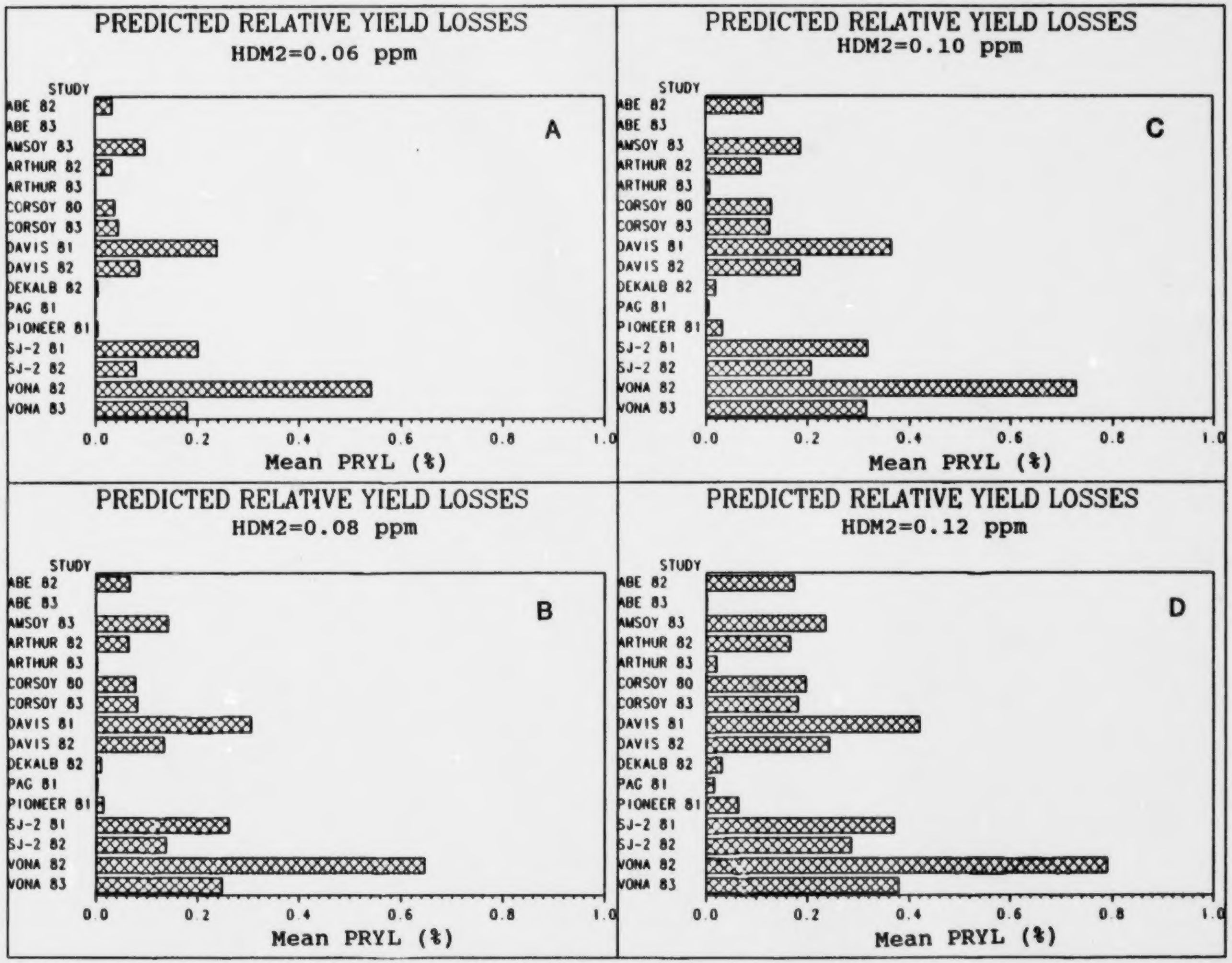


Figure 3.2. Predicted relative yield losses (PRYLs) for 16 NCLAN studies using the Weibull model and HDM2 at (A) 0.06 ppm, (B) 0.08 ppm, (C) 0.10 ppm, and (D) 0.12 ppm.

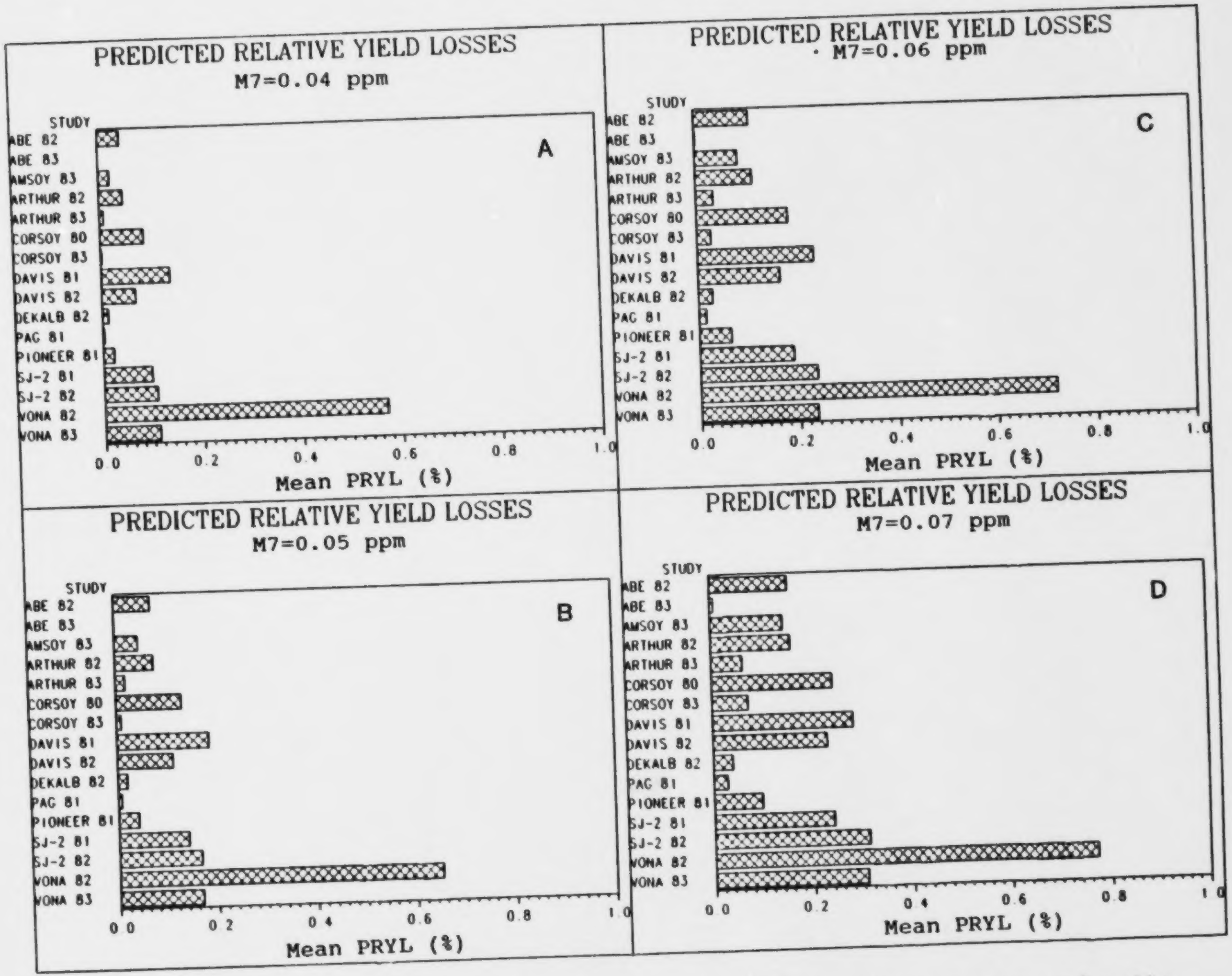


Figure 3.3. Predicted relative yield losses (PRYLs) for 16 NCLAN studies using the Weibull model and M7 at (A) 0.04 ppm, (B) 0.05 ppm, (C) 0.06 ppm, and (D) 0.07 ppm.

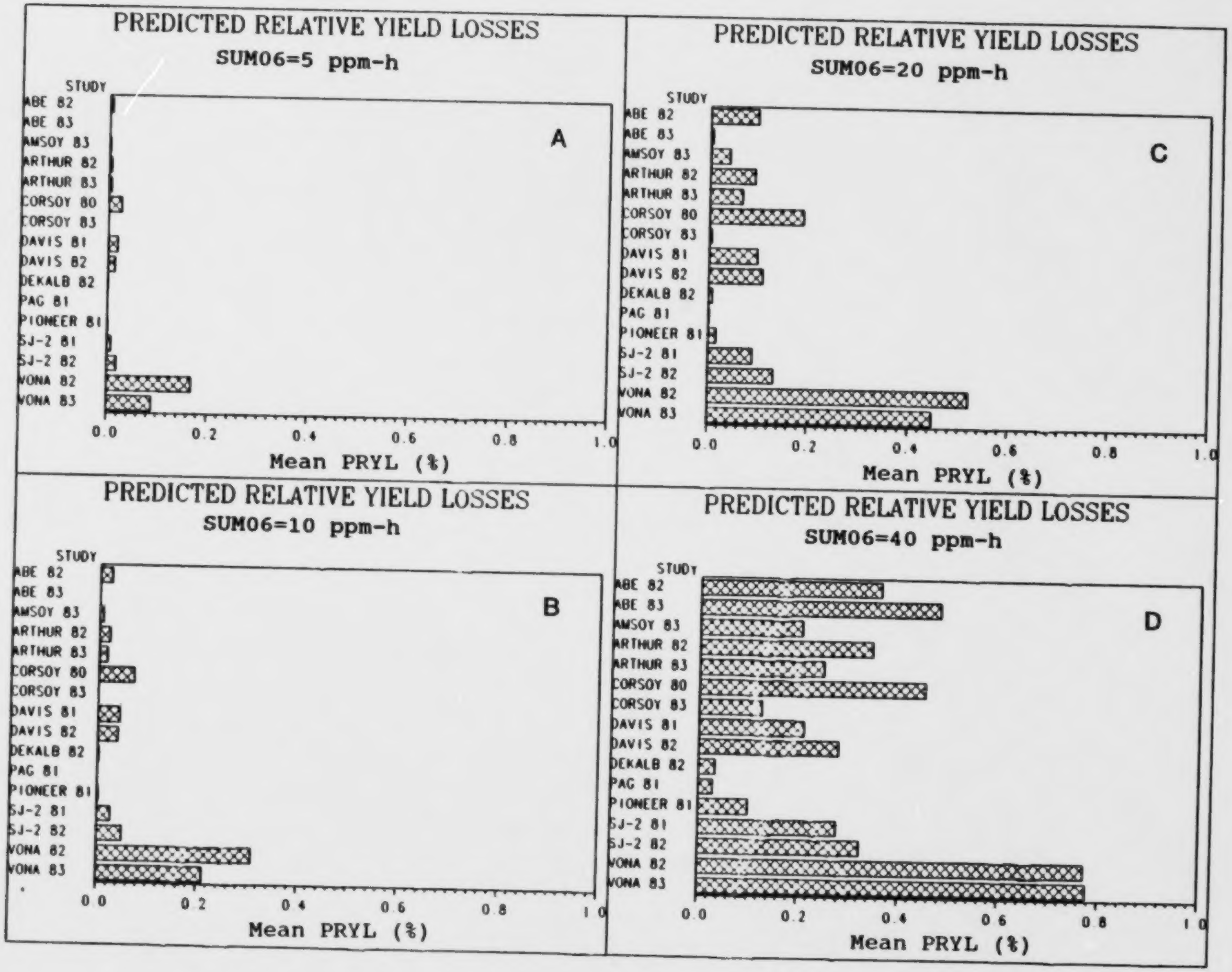


Figure 3.4. Predicted relative yield losses (PRYLs) for 16 NCLAN studies using the Weibull model and SUM06 at (A) 5 ppm-h, (B) 10 ppm-h, (C) 20 ppm-h, and (D) 40 ppm-h.

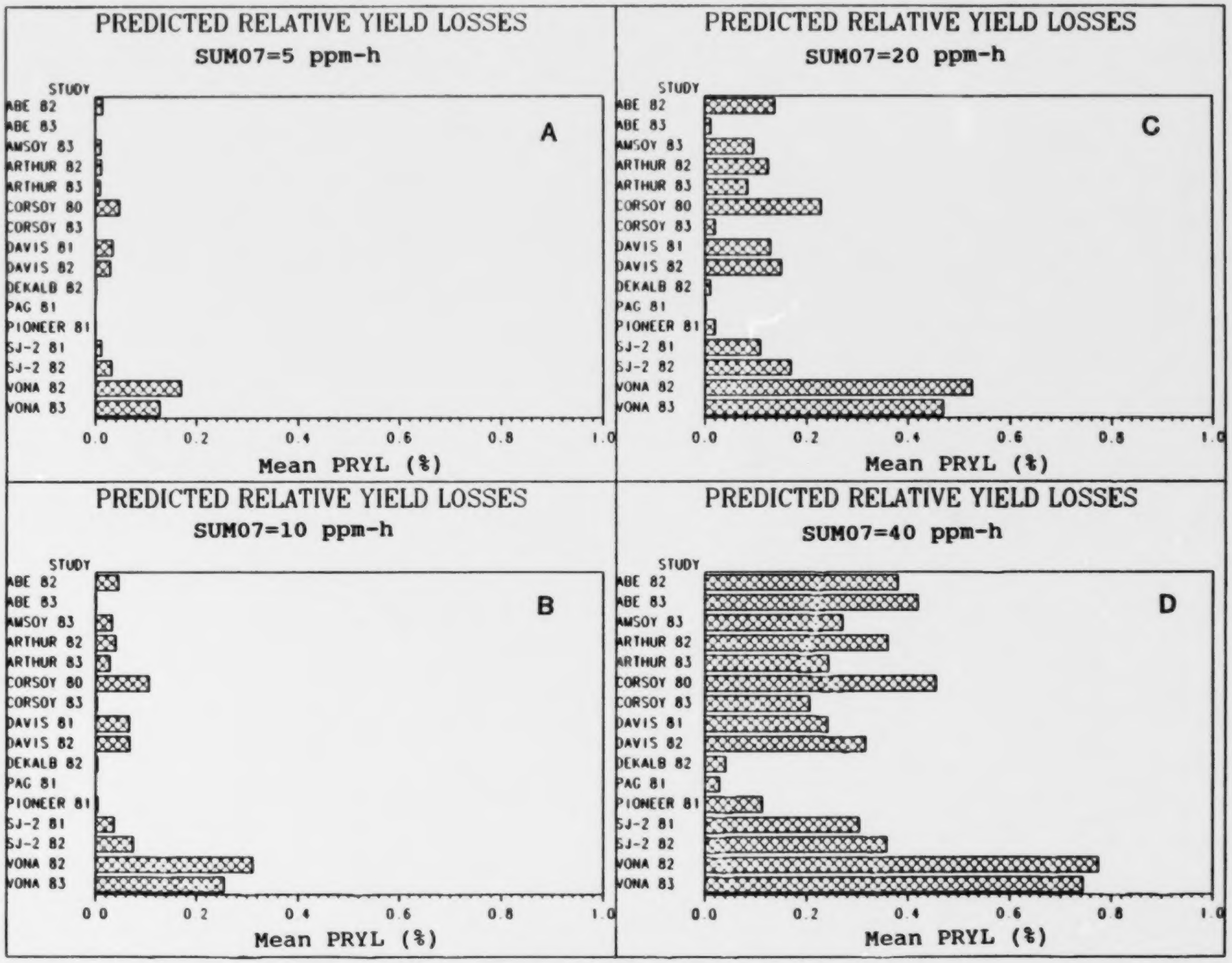


Figure 3.5. Predicted relative yield losses (PRYLs) for 16 NCLAN studies using the Weibull model and SUM07 at (A) 5 ppm-h, (B) 10 ppm-h, (C) 20 ppm-h, and (D) 40 ppm-h.

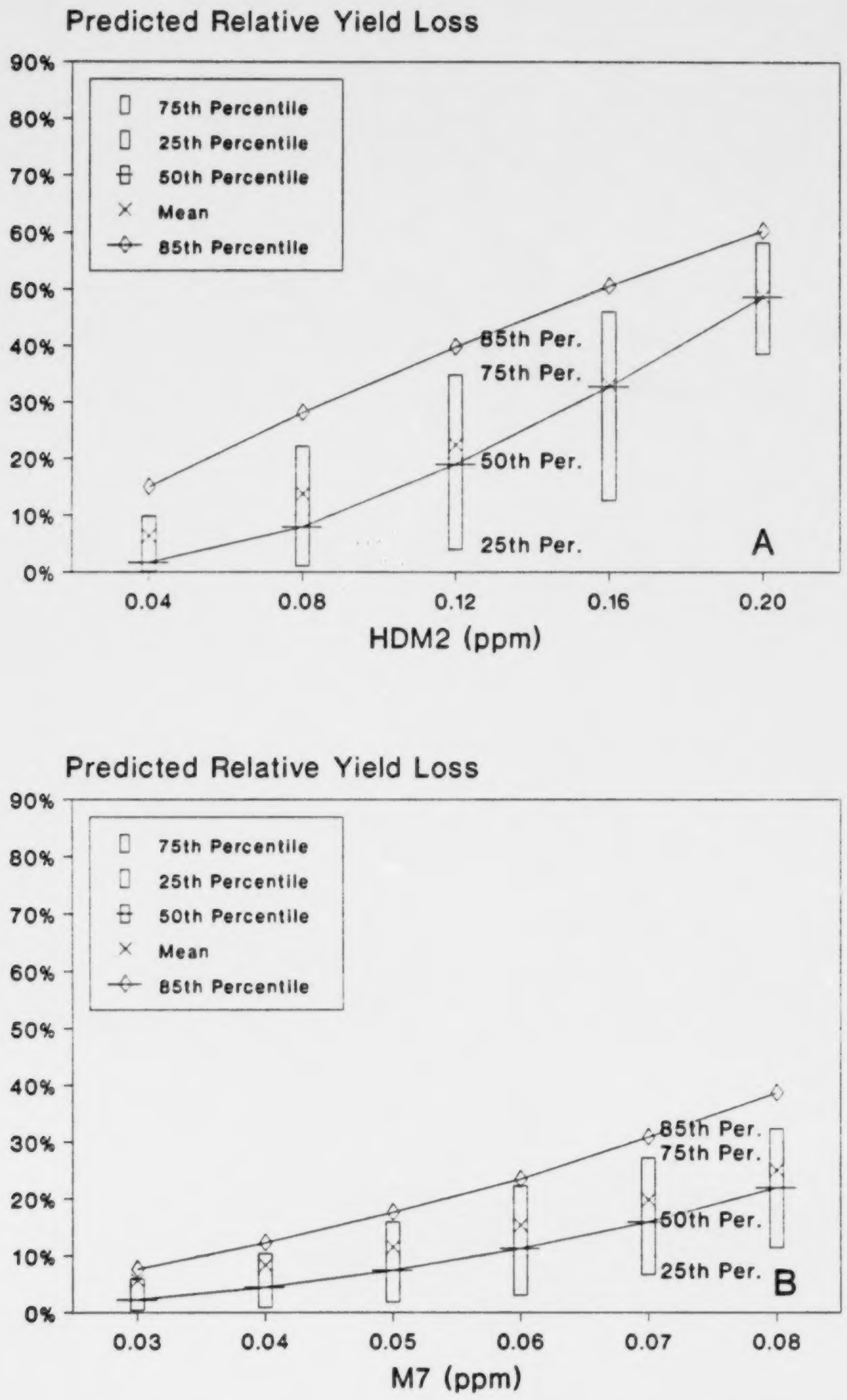


Figure 3.6. Boxplots of percentiles and mean predicted relative yield losses, characterizing the range of crop losses across the 16 NCLAN studies using the Weibull model with (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07.

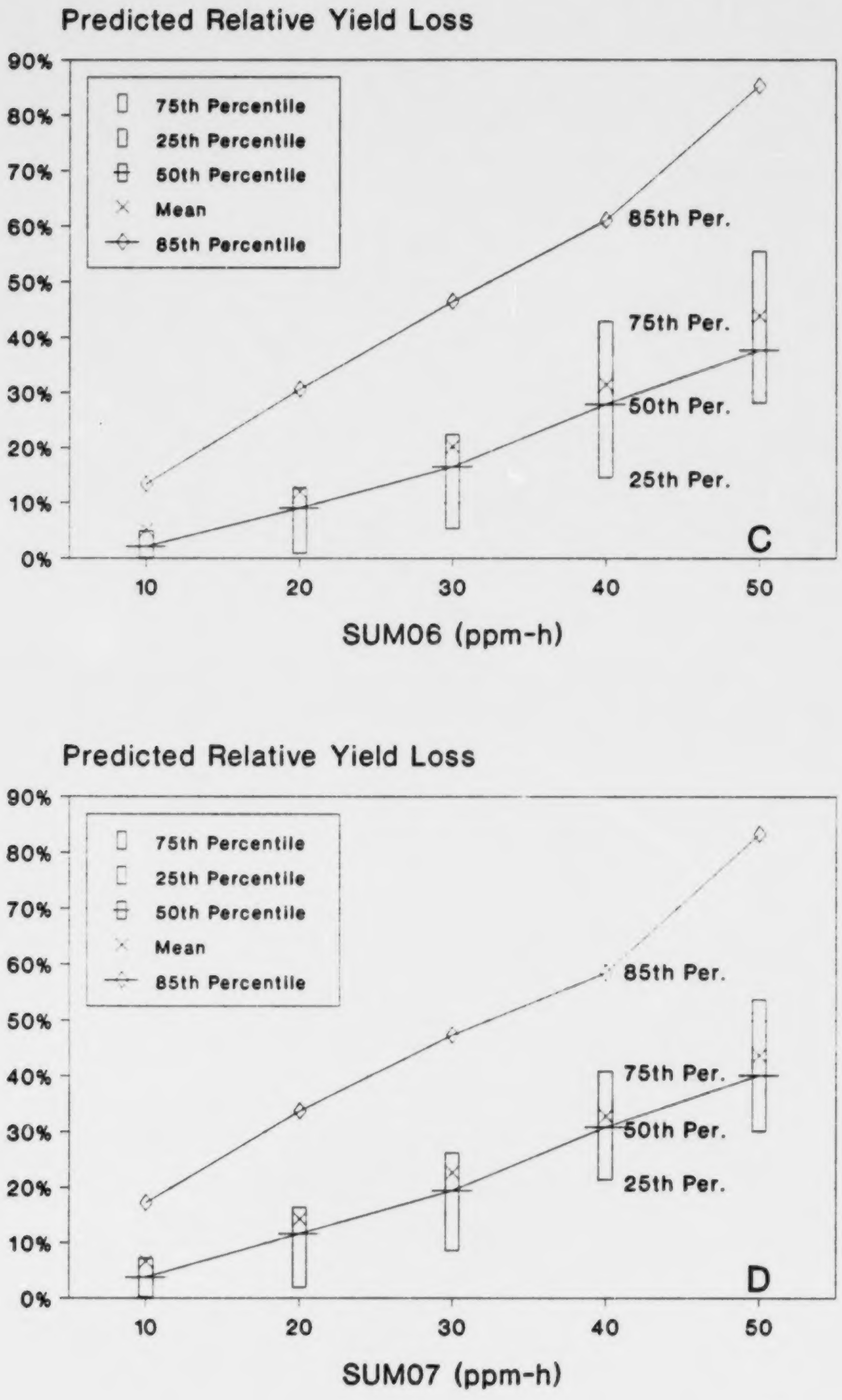


Figure 3.6. (cont)

VL

CHAPTER 4

RELATIONSHIPS AMONG EXPOSURE INDICES AT 83 RURAL AND AGRICULTURAL SITES

4.1. INTRODUCTION

The scientific and technical information on O_3 and its effects on vegetation, natural ecosystems, and human welfare indicate that a secondary ambient air quality standard is needed that correlates well with both short- and long-term exposures (U.S. EPA, 1987). Key issues of concern for a secondary ambient air quality standard are the relationships of various air quality indicators and aggregation periods.

The search for a single, secondary ambient air quality standard that protects against repeated peaks and long-term exposures has proceeded along two lines: (1) air quality analysis or correlation analysis of ambient air quality data, and (2) regression analysis of biological response against exposure indices. Air quality analyses focus on identifying the relationships among exposure indices and selecting indices that correlate well with various exposure patterns. In particular, an appropriate air quality indicator should correlate well with short-term peak, multiple peak, and long-term average indicators (U.S. EPA, 1987).

Air quality analysis of 86 non-urban areas found that the maximum monthly mean of 1-h daily maximum O_3 concentrations correlated well with short- and long-term exposure indices (U.S. EPA, 1987). This air quality indicator has been proposed as a surrogate of both short- and long-term exposure indices. But, Lee et al. (1988a) found that while the "seasonal" exposure indices adequately depicted the plant response to exposure, the monthly exposure indices did not relate well to crop yield loss.

Previous studies using both experimental and ambient air quality data report that the relationships between peak and mean indices are unstable (Johnson et al., 1986, Heck et al., 1984a, 1984b; Larsen and Heck, 1985). Relationships between exposure indices using NCLAN and SAROAD data show considerable scatter among the three groups of exposure indices: Mean (e.g. M7, M1), One Event (e.g., HDM2, P1, P7), and Concentration Weighting (e.g., HRS08, HRS10). Correlations among indices from different groups are no better than 0.80.

The primary objective in this chapter is to determine the relationships among the current secondary NAAQS (i.e., HDM2, a peak or one-event index), the 7-hour seasonal mean (M7), a long-term average that NCLAN adopted in exposure-response modeling, and the two cumulative censored indices (i.e., SUM06 and SUM07) that related well to plant production. We extended the air quality analyses of a previous study (U.S. EPA, 1987) by developing regression equations that relate one index to another using ambient air monitoring data. A secondary objective is to determine the aggregation period for describing the ambient exposures for agricultural/remote sites using SUM06 and SUM07.

4.2. METHOD

4.2.1. AMBIENT OZONE MONITORING DATA

Air quality data obtained from the National Air Data Branch are representative of 83 "non-urban" monitoring sites with a valid daily maximum 1-hour concentration for at least 75% of the days in an EPA-specified O_3 season. These sites were classified according to local land use as "rural/agricultural", "strictly remote", or "other remote" by Johnson et al. (1986) and appeared in previous air pollution studies (i.e., U.S. EPA, 1987; Larsen et al., 1989). The data cover the 1982-1986 time period and represent 34 states (see Table 4.1).

4.2.2. OZONE EXPOSURE INDICES

In this section, we restrict our attention to the current secondary NAAQS for O_3 (i.e., the second highest daily maximum 1-h concentration, HDM2), the 7-h seasonal mean concentration for 0900-1600 h (M7), and the integration of all hourly concentrations of 0.06 (0.07) ppm and higher (SUM06, SUM07). In calculating the four numerical indices for the 83 site-years of O_3 data, missing values were not filled in. The HDM2 index was calculated using all available hourly concentrations and it was determined which sites attain (i.e., HDM2≤0.12 ppm) or exceed (i.e., HDM2>0.12 ppm) the level of the current secondary NAAQS. The other three indices were calculated using hourly concentrations for the five months, May to September, which correspond to the growing season for most crops in the United States.

Three-, four-, and five-month periods from May to September have been considered in attempts to determine the time period that captures the majority of hourly concentrations of 0.06 ppm or higher. The results, in Table 4.2, indicate the five-month period from May to September should be used to capture approximately 90% of the peak concentrations. For most states, the May to September time period contains the EPA-specified O_3 season.

Further, in 73 of the 83 site-years, the HDM2 occurs within the 5-month period. Another five site-years have the HDM2 occurring in the last week of April.

48

In the definition of the censored indices, we chose to include the cutoff level in the calculation. For example, SUM06 cumulates hourly concentrations of 0.06 ppm or higher over time (see Figure 2.1). Previous studies (i.e., Lee et al., 1987; Johnson et al., 1986; U.S. EPA, 1987) defined the cumulative censored indices excluding the cutoff level. Theoretically, O_3 concentrations are continuously distributed and point probabilities are zero. However, ambient O_3 monitoring data are discrete, and, in many cases, recorded to the nearest 0.005 or 0.01 ppm. For these sites, the inclusion or exclusion of the cutoff level can affect the calculation by as much as 200%.

4.2.3. REGRESSION ANALYSIS AND JOINT DISTRIBUTIONS

Relationships among exposure indices were analyzed using polynomial regression models and joint distributions among pairs of indices (i.e., frequencies of co-occurrence). Marginal distributions (i.e., histograms) of the four exposure indices were used to characterize the agricultural/remote SAROAD sites and determine the range of ambient conditions for these sites. A test for distributional differences for M7, SUM06, and SUM07 between sites with $HDM2 \leq 0.12$ ppm and those with $HDM2 > 0.12$ ppm was performed using the Wilcoxon test for two independent samples.

Simple linear and quadratic models were used to approximate the true unknown relationships among the four exposure indices. These regression equations were used to relate the exposure indices to determine values of one index in terms of another index.

Finally, the joint distribution of exposure indices (two at a time) were considered to determine if sites dichotimized according to one index could be consistently grouped according to a second index. The joint distribution is useful for relating exposure indices by determining the probabilities of co-occurrence. For given values of one index, the conditional probabilities give the range of possible values of a second index.

4.3. RESULTS

4.3.1. EMPIRICAL DISTRIBUTIONS OF EXPOSURE INDICES

Hourly distributions of the 83 site-years were characterized by the four exposure indices (Table 4.1). Two sites, San Bernardino County, CA and Knox County, TN, had extremely high O_3 readings for at least one index. For San Bernardino, the SUM06 value of 167.9 ppm-h was 50% higher than the next highest SUM06 value of 111.7 for Kern County, CA. A similar pattern was observed for SUM07. For Knox County, the peak index, HDM2, was 0.240 ppm, 20% above the next highest HDM2 value. In contrast, the mean and cumulative indices for Knox County were average.

(43)

Based on the HDM2 index, 33 of the 83 sites (40%) had HDM2>0.12 ppm. For another 27 (33%) sites, the concentration of HDM2 was between 0.11 and 0.12 ppm. The empirical distribution of HDM2 values for the 83 site-years (Figure 4.1A), showed the average HDM2 value to be 0.123 ppm. Table 4.3 gives percentiles and other summary statistics of the empirical distributions, which can be interpolated using the cumulative distribution functions (CDF) and the boxplots (below the X-axis) in Figure 4.1. The vertical lines of the boxplot represent the 25th, 50th, and 75th percentiles; the mean is denoted by an asterisk.

Depending upon the measure of exposure, the ambient air quality data for the 83 agricultural/remote site-years displayed small to large differences in exposure patterns. The M7 index indicated that the O_3 exposures for the 83 site-years were fairly uniform, with 90% of the site-years having M7 values between 0.03 and 0.06 ppm. The M7 index was the least variable with a coefficient of variation (CV) of 19% (Table 4.3). That is, about 66% of the population lies within 19% of the mean, based upon M7.

HDM2 values ranged from 0.060 to 0.240 ppm for the 83 site-years of ambient data, with an average of 0.123 and coefficient of variation of 25%. Over 50% of the agricultural/remote site-years were within ±0.02 ppm of the level (0.12 ppm) of the current standard.

Greater differences in exposure patterns were indicated by the cumulative censored indices, SUM06 and SUM07, than by the single peak or mean indices. The population coefficients of variation for SUM06 and SUM07 were 66% and 87%, respectively; or 58% and 77%, respectively, excluding San Bernardino County, CA which had the highest SUM06 and SUM07 values. The integrated indices suggested that a broad range of exposure patterns existed among the 83 site-years. Figures 4.1C-D show that the empirical distributions of SUM06 and SUM07 are highly skewed with an extended right tail. For SUM06, about 66% of the population lies between 14 and 64 ppm-h. For SUM07, about 66% of the population lies between 3 and 44 ppm-h.

Excluding the San Bernardino County, CA site with extreme values of SUM06 and SUM07, the two independent-sample Wilcoxon test was used to compare the mean exposures for sites at which the concentration of HDM2 was ≤0.12 ppm and those for which HDM2 was >0.12 ppm. The means, standard deviations, and P-values (one-sided test that hypothesizes a greater mean exposure index for sites not attaining the current standard than for sites attaining the standard) were:

50

	HDM2≤0.12 ppm	HDM2>0.12 ppm	Wilcoxon test P-value
M7	0.043 (0.008)	0.050 (0.008)	0.023
SUM06	29.9 (18.1)	48.9 (21.4)	0.017
SUM07	15.8 (11.7)	32.6 (19.1)	0.010

Significant differences between the two groups of sites (HDM2≤0.12 and HDM2>0.12 ppm) were noted for the three exposure indices. These differences in exposure patterns were clearer for SUM06 and SUM07 than for M7, i.e., sites with HDM2≤0.12 and sites with HDM2>0.12 ppm differed by 64% and 106%, based on SUM06 and SUM07, respectively, versus 16% for M7. Sites with HDM2>0.12 ppm had considerably more concentrations of 0.06 (or 0.07) ppm or higher than sites with HDM2≤0.12 ppm. This was not revealed in the M7 index.

4.3.2. RELATIONSHIPS AMONG EXPOSURE INDICES

Figure 4.2 shows the scatterplots of all possible pairs of exposure indices and the polynomial curves from the regressions. The two extreme sites, San Bernardino, CA and Knox, TN, are highlighted by asterisks. These points had large influence and/or large residuals in the regression calculations, often forcing a quadratic term to be included in the model. In some cases, the quadratic term was negative, implying an inverse relationship between exposure indices beyond some level. For these reasons, the two sites were excluded from the regression calculations.

Table 4.4 gives the polynomial models that were fit to the remaining 81 SAROAD sites. The R-squared (RSQ) is the proportion of variation of the dependent variable (Y) explained by the independent variable (X). For the simple linear model, the square root of RSQ is the Pearson correlation coefficient between X and Y. The root mean square error (ROOT MSE) estimates the standard deviation about the regression line.

A linear model was adequate for all pairs, except for SUM06 and SUM07 against M7, which required quadratic models. As noted in previous air quality analyses (Heck et al., 1984a; Cure et al., 1986; Johnson et al., 1986; U.S. EPA, 1987), the relationship between the mean and peak indices was poor (RSQ=0.288). Better relationships existed between the censored cumulative indices, SUM06 and SUM07, and the peak index, HDM2 (RSQ=0.451 and 0.509, respectively); and the mean index, M7 (RSQ=0.755 for both).

A correlation of 0.75 (or a RSQ of 0.562) denoted a strong association between air quality indicators (U.S. EPA, 1987). The integrated indices, SUM06 and SUM07, were strongly related to M7, and less related to HDM2. However, the relationship between

51

SUM07 and HDM2 (RSQ=0.509; r=0.71) fell just below the level defined for a strong association.

Based on the regression equations involving HDM2, comparable values of the other three exposure indices can be calculated. In particular, HDM2 values of 0.10 ppm and 0.12 ppm were used to obtain comparable values for the other three indices in one of two ways: (1) inversion of the linear models regressing HDM2 on the other index, and (2) interpolation of the linear models regressing the other index on HDM2. These values are:

	HDM2 with (1)		HDM2 with (2)	
VAR	0.10	0.12	0.10	0.12
SUM06	13.4	36.5	26.6	36.9
SUM07	4.6	21.8	13.4	22.1
M7	0.034	0.045	0.042	0.045

For HDM2 equal to 0.12 ppm, the two methods gave similar values for M7, SUM06, and SUM07. However, for HDM2 equal to 0.10 ppm, the two methods gave considerably different values for the three indices.

In mathematics, the equation to calculate X_1 given X_2 is identical to the equation to calculate X_2 given X_1. In statistics, the regression equation of X_1 on X_2 differs from the regression equation of X_2 on X_1, due to data variation and model uncertainty. Each equation has a different interpretation and use. For any X_1, there is a population of values of X_2 with some distribution having conditional mean $u_{2.1}$. The population regression of X_2 on X_1 is the relationship of the means $u_{2.1}$ to X_1. In this sense, for any X_1, the comparable value of X_2 is $u_{2.1}$ (e.g., Snedecor and Cochran, 1980). Vice versa, for any X_2, the comparable value of X_1 is $u_{1.2}$.

For any HDM2, the comparable values of M7, SUM06, and SUM07 were derived from the regressions of the exposure indices on HDM2. That is, method (2) was the more appropriate calculation.

For a 17% decrease in HDM2 (from 0.12 to 0.10 ppm), one might expect a 28% decrease in SUM06, a 39% decrease in SUM07, and a 7% decrease in M7. The M7 index was insensitive to incremental changes in the peak index while SUM06 and SUM07 were very sensitive to changes in HDM2. Since the mean weights all concentrations equally, the M7 index was relatively insensitive to changes in the peak concentrations. On the other hand, changes in the peak index translated to greater changes in the cumulative censored indices.

4.3.3. JOINT DISTRIBUTIONS OF EXPOSURE INDICES

The scatterplot matrix in Figure 4.2 shows fair to strong correlations among the exposure indices. However, the density of points in the scatterplots was difficult to comprehend and did not reveal the frequency of co-occurrence among pairs of exposure indices. This was especially true in the middle of the data cluster where the high density of points obscured the relationship among indices. This area was of particular interest for determining the range of values of M7, SUM06, and SUM07, associated with values of HDM2 around the current standard of 0.12 ppm.

The joint distributions of the exposure indices were given in Tables 4.5A-F and illustrated in Figures 4.3A-F. All 83 site-years were included in this analysis. The last row and column of Tables 4.5A-F are the marginal distributions (i.e., histograms), scaled by the sample size. The remaining rows and columns are the conditional distributions, scaled by the marginal total.

For fixed values of M7, a wide range of values for SUM06, SUM07, and HDM2 are possible (Figures 4.3A,D,E). For example, the site-years with M7 between 0.042 and 0.048 ppm were associated with HDM2 values from 0.06 to 0.24 ppm (Table 4.5A), SUM06 values from 0 to 70 ppm-h (Table 4.5D), and SUM07 values from 0 to 40 (Table 4.5E). Many different exposure patterns with wide ranging values for HDM2, SUM06, and SUM07 can and do occur, for values of M7 between 0.042 and 0.048 ppm. The M7 index did not distinguish among these patterns, although some were more likely to occur than others.

Table 4.5A showed that 9 of the 20 site-years with average (M7) O_3 concentrations between 0.042 and 0.048 ppm had HDM2 concentrations $\leq$0.12 ppm while the concentration at the other 11 sites exceeded 0.12 ppm. On the other hand, SUM06 and SUM07 show a much stronger relationship. The joint distribution of SUM06 and SUM07 has non-zero cell frequencies in a tight band along the main diagonal (Table 4.5F). The joint distributions of other pairs lie somewhere between these two extreme patterns.

There were 51 sites with HDM2 between 0.10 and 0.14 ppm; of these 33 sites had HDM2 concentrations $\leq$0.12 ppm, while at the other 18 sites the concentrations of the HDM2 exceeded 0.12 ppm. These sites have M7 values ranging from 0.024 to 0.066 ppm (Table 4.5A), SUM06 values ranging from 0 to 84 ppm-h (Table 4.5B), and SUM07 values ranging from 0 to 60 ppm-h (Table 4.5C), with varying probabilities. These results indicate that the values of M7, SUM06, and SUM07 cannot be predicted with much precision given values of HDM2 between 0.10 and 0.14 ppm.

Sites just below and above the level (0.12 ppm) of the current standard have similar conditional distributions of M7, SUM06, and SUM07. That is, the values of M7, SUM06, and SUM07 are equally unpredictable for sites with HDM2 between 0.10 and 0.12 ppm and sites with HDM2 between 0.12 and 0.14 ppm. These results indicate that an exposure level using HDM2 does not readily translate into an exposure level using any of the other three air quality indicators.

4.4. DISCUSSION

The relationships among the four exposure indices (HDM2, M7, SUM06, SUM07) were consistent with the relationships found in previous studies using similar indices (Heck *et al*., 1984a, 1984b; Johnson *et al*., 1986; U.S. EPA, 1987; Larsen *et al*., 1989). Mainly, the relationship among peak and mean air quality indicators was weak. High (low) values of HDM2 were associated with high (low) values of M7 but considerable scatter existed in the mid-values.

A criterion for an ambient air quality standard is that the air quality indicator must correlate well with both short- and long-term indicators (U.S. EPA, 1987). The cumulative censored indices, SUM06 and SUM07, had strong relationships with M7 but only good relationships with HDM2. The association between the integrated indices and the peak or mean indices was strong at the extremes and less predictable in the middle.

Exposure patterns for the 83 non-urban site-years had similar long-term averages but differed widely in temporal variation. This was most evident in the SUM06 (SUM07) air quality index, defined for the five-month period from May to September, which ranged from 1 to 168 ppm-h. Some sites had few hourly concentrations of 0.06 ppm or higher while other sites had numerous concentrations of 0.06 ppm or higher.

The seasonal 7-h mean O_3 weighted all concentrations equally and did not adequately reflect the temporal variations in the exposures. An infinite number of different exposure patterns can have the same mean concentration. Sites with M7 values around 0.045 ppm were associated with HDM2 values from 0.06 to 0.24 ppm and SUM06 values from 1 to 70 ppm-h.

4.5. CONCLUSION

An aggregation period of 5 months (May to September) was used to calculate air quality indicators for 83 non-urban sites. The five-month period captured about 90% of the hourly concentrations of 0.06 ppm or higher, included the EPA-defined O_3 season for most sites, represented the period when the second highest daily maximum 1-h concentration occurs, and coincided with the agricultural season for many crops in the United States.

An air quality analysis of 83 non-urban site-years showed that the cumulative censored indices, SUM06 and SUM07, correlated well with a long-term index (M7) and a short-term index (HDM2). Both indices were sensitive to short- and long-term exposures and related well to biological response. These results suggested that SUM06 and SUM07 had potential for setting a standard that protects against adverse effects from repeated peak and long-term exposures.

TABLE 4.1.

Rural agricultural and remote monitoring sites reporting ozone data from SAROAD.

SAROAD ID	ST	COUNTY	YEAR	M7 (ppm)	SUM06 (ppm-h)	SUM07 (ppm-h)	HDM2 (ppm)
030050110A08	AZ	Greenlee Co	82	0.046	11.8	0.4	0.075
051620002F01	CA	Contra Costa Co	82	0.040	20.6	11.8	0.110
052820002I01	CA	Fresno Co	83	0.048	28.3	13.3	0.100
052820003I01	CA	Fresno Co	82	0.046	62.9	35.9	0.110
053480007F01	CA	Kern Co	86	0.074	111.7	89.2	0.160
055560001I01	CA	Ventura Co	86	0.053	47.4	26.2	0.170
056405810I01	CA	Placer Co	85	0.050	49.2	33.4	0.160
057200008I01	CA	Santa Barbara Co	84	0.046	28.7	13.8	0.140
057220001I01	CA	Santa Barbara Co	83	0.048	28.7	11.7	0.140
058390001I01	CA	San Bernardino Co	83	0.066	167.9	128.2	0.160
058900002I01	CA	Sutter Co	82	0.045	39.9	23.5	0.100
071060001F01	CT	Tolland Co	83	0.060	87.3	63.5	0.190
080180018F01	DE	New Castle Co	83	0.055	78.8	62.0	0.170
102340002F01	FL	Leon Co	85	0.030	3.8	0.7	0.100
142220001F03	IL	Effingham Co	84	0.054	47.0	27.1	0.110
144640002F03	IL	Macoupin Co	84	0.051	40.8	21.6	0.130
144680007F01	IL	Madison Co	83	0.058	57.6	41.3	0.153
147960001F03	IL	Monroe Co	84	0.053	43.4	23.7	0.138
148320007F03	IL	Will Co	83	0.046	29.8	15.0	0.124
152040042H03	IN	Marion Co	85	0.047	43.4	21.7	0.119
153420021G01	IN	Porter Co	82	0.037	27.1	12.9	0.132
153700007G01	IN	St Joseph Co	83	0.054	61.1	44.9	0.135
162360003G01	IA	Madison Co	83	0.040	29.6	9.6	0.095
163120024G01	IA	Polk Co	83	0.052	66.9	42.4	0.113
171720001F05	KS	Jefferson Co	83	0.035	20.1	2.0	0.110
181360001F05	KY	Fulton Co	83	0.050	50.0	23.4	0.101
181600004F05	KY	Hardin Co	83	0.040	44.1	25.5	0.113
183460002F05	KY	Pulaski Co	83	0.037	28.1	13.7	0.107
190500001F01	LA	Caddo Parrish	85	0.048	29.7	15.3	0.114
201205016F05	ME	Washington Co	84	0.036	21.6	12.7	0.115
210080014G01	MD	Anne Arundel Co	84	0.046	52.2	35.9	0.151
210360001F01	MD	Carroll Co	84	0.050	36.7	22.4	0.139
210740034G01	MD	Frederick Co	83	0.054	53.1	38.3	0.119
220030003F01	MA	Hampden Co	83	0.050	65.9	50.5	0.185
220060002F01	MA	Hampshire Co	84	0.041	28.8	18.6	0.125
222196001F01	MA	Middlesex Co	84	0.047	36.1	24.5	0.147
231040001F02	MI	Clinton Co	86	0.037	21.3	9.2	0.103
231700001F02	MI	Genesee County	86	0.036	13.2	4.2	0.110
232720001H02	MI	Kent Co	83	0.045	44.0	27.4	0.127
244000027F01	MN	Washington Co	86	0.035	8.6	3.3	0.096
262950001A08	MO	Madison Co	83	0.049	74.5	40.3	0.110
270310101A08	MT	Rosebud Co	83	0.042	4.2	0.0	0.065
311500001F01	NJ	Hunterdon Co	82	0.051	53.6	39.4	0.196
313260001F01	NJ	Morris Co	83	0.057	69.6	52.4	0.185
320340008F02	NM	Dona Anna Co	86	0.047	39.0	17.7	0.110
332000005F01	NY	Erie Co	83	0.044	35.5	19.4	0.116
332020002F03	NY	Essex Co	83	0.043	39.9	22.3	0.110
332020005F03	NY	Essex Co	84	0.037	24.4	12.6	0.121
334380005F01	NY	Monroe Co	82	0.028	13.1	5.7	0.094
334720006F01	NY	Niagara Co	86	0.039	27.8	17.2	0.110
335060004F01	NY	Oneida Co	83	0.038	21.4	10.9	0.110
337240001F01	NY	Wayne Co	83	0.044	42.1	26.8	0.121
340945101A08	NC	Johnston Co	82	0.039	11.0	4.7	0.090
342120099F05	NC	Jones Co	84	0.043	31.0	16.7	0.098
342560099F05	NC	Martin Co	84	0.050	34.0	19.3	0.100
342660099F05	NC	Montgomery Co	83	0.051	52.7	28.3	0.115
342880002F03	NC	New Hanover Co	82	0.039	18.3	5.9	0.099
343140099F05	NC	Person Co	86	0.054	52.8	30.4	0.108
344180001F03	NC	Wake Co	83	0.053	49.0	33.9	0.133
350340003F03	ND	Dunn Co	83	0.043	4.3	0.5	0.072
361340001F03	OH	Clinton Co	86	0.046	38.2	22.3	0.116
371740073F03	OK	Mcclain Co	85	0.041	28.9	19.7	0.107
373020174F03	OK	Tulsa Co	85	0.050	33.9	16.2	0.122
380260004F03	OR	Clackamas Co	85	0.032	16.6	9.1	0.130
380300004F03	OR	Columbia Co	82	0.028	6.7	3.1	0.120
380420111A08	OR	Crook Co	83	0.038	1.1	0.0	0.060
381140004F01	OR	Marion Co	85	0.030	9.6	4.5	0.110
397120301F03	PA	Perry Co	82	0.040	22.9	13.3	0.130
420980001F03	SC	Edgefield Co	83	0.054	51.3	31.1	0.117
421200007F01	SC	Greenville Co	83	0.053	51.2	33.3	0.113
421900001F03	SC	Richland Co	85	0.049	37.8	21.8	0.110
441280011G03	TN	Hamilton Co	83	0.053	58.6	40.3	0.200
441720021G03	TN	Knox Co	86	0.043	38.3	17.2	0.240
451420054H01	TX	Denton Co	84	0.060	85.6	64.7	0.160
452180001F03	TX	Gregg Co	82	0.047	42.1	24.4	0.130
460900004F05	UT	Salt Lake Co	82	0.051	24.6	12.1	0.115
481120002F03	VA	Fauquier Co	85	0.053	49.9	28.4	0.112
481460001F01	VA	Hanover Co	83	0.061	75.3	56.0	0.140
482180001F01	VA	Northampton Co	84	0.051	42.0	22.2	0.113
491560005F01	WA	Pierce Co	82	0.030	11.7	7.3	0.100
492000001F01	WA	Snohomish Co	85	0.032	9.7	4.0	0.110
510490001A08	WI	Taylor Co	82	0.035	12.5	3.1	0.075
511650007F03	WI	Marathon Co	85	0.041	15.6	5.0	0.087

(56)

TABLE 4.2.

Proportion of the annual SUM06 and SUM07 values that occur within a 3-, 4-, or 5-month period from May to September.

	May to July	June to Aug	July to Sept	May to Aug	June to Sept	May to Sept
SUM06	58%	63%	52%	78%	72%	88%
SUM07	59%	66%	55%	80%	75%	89%

Note: The maximum 3-month period between May and September captures 69 and 73% of the annual SUM06 and SUM07 values, respectively. The maximum 4-month period between May and September captures 81 and 83% of the annual SUM06 and SUM07 values, respectively.

TABLE 4.3.

Cumulative distribution functions (CDF) of the four exposure indices, HDM2, SUM06, SUM07, and M7 for the 83 SAROAD sites.

CDF	HDM2 (ppm)	M7 (ppm)	SUM06 (ppm-h)	SUM07 (ppm-h)
0%	0.060	0.028	1.1	0.0
1%	0.060	0.028	1.1	0.0
5%	0.075	0.030	4.8	0.5
10%	0.094	0.035	10.2	3.2
25%	0.107	0.039	21.4	10.9
50%	0.115	0.046	36.7	19.6
75%	0.135	0.051	51.3	31.1
90%	0.166	0.055	68.5	48.2
95%	0.189	0.060	84.2	63.2
99%	0.240	0.074	167.9	128.2
100%	0.240	0.074	167.9	128.2
Mean	0.123	0.046	38.9	23.7
Std Dev	0.031	0.009	25.9	20.7
CV	25%	19%	66%	87%

TABLE 4.4.

Relationships among exposure indices, HDM2, SUM06, SUM07, and M7.

DEP VAR	INDEP VAR	INT	LIN	QUAD	RSQ	ROOT MSE	MEAN	CV
HDM2	M7	0.040	1.770		0.288	0.0239	0.121	20%
HDM2	SUM06	0.088	0.0009		0.451	0.0210	0.121	17%
HDM2	SUM07	0.095	0.0012		0.509	0.0199	0.121	16%
M7	HDM2	0.026	0.163		0.288	0.0073	0.046	16%
M7	SUM06	0.033	0.0003		0.733	0.0044	0.046	10%
M7	SUM07	0.036	0.0004		0.689	0.0048	0.046	11%
SUM06	M7	-13.9	0.0000	23852	0.755	10.87	37.3	29%
SUM06	HDM2	-25.4	519.6		0.451	16.28	37.3	44%
SUM06	SUM07	9.7	1.229		0.952	4.81	37.3	13%
SUM07	M7	29.2	-2066.4	40699	0.755	8.68	22.4	39%
SUM07	HDM2	-30.5	438.4		0.509	12.22	22.4	54%
SUM07	SUM06	-6.5	0.775		0.952	3.82	22.4	17%

HDM2 and M7 were reported in ppm. SUM06 and SUM07 were reported in ppm-h. The intercept (INT), root MSE, and MEAN were reported in the same units as the dependent variable. Two sites, San Bernardino County, CA (1983) and Knox County, TN (1986), were excluded from the regression calculations due to extremely large ozone levels for SUM06 and HDM2, respectively. The large SUM06 value of 167.9 ppm-h for San Bernardino County, CA had high influence on the regression coefficients, requiring a quadratic model to fit the extreme point well. The large HDM2 value of 0.24 ppm for Knox County, TN was poorly fit using the polynomial models and accounted for 30% of the total unexplained variation in a number of the regressions.

TABLE 4.5.

Joint distributions of the four exposure indices HDM2, M7, SUM06, and SUM07.

A) M7 (ppm)

HDM2 \ LOWER	0.024	0.030	0.036	0.042	0.048	0.054	0.060	0.066	TOTAL
(ppm) \ UPPER	0.030	0.036	0.042	0.048	0.054	0.060	0.066	0.074	
(0.04, 0.06]	0	0	1	0	0	0	0	0	1
(0.06, 0.08]	0	1	0	3	0	0	0	0	4
(0.08, 0.10]	3	1	4	2	2	0	0	0	12
(0.10, 0.12]	1	4	8	6	11	3	0	0	33
(0.12, 0.14]	0	1	4	6	5	1	1	0	18
(0.14, 0.16]	0	0	0	2	1	2	1	1	7
(0.16, 0.18]	0	0	0	0	1	1	0	0	2
(0.18, 0.24]	0	0	0	1	3	1	1	0	6
TOTAL	4	7	17	20	23	8	3	1	83

B) SUM06 (ppm-h)

HDM2 \ LOWER	0	14	28	42	56	70	84	98	TOTAL
(ppm) \ UPPER	14	28	42	56	70	84	98	168	
(0.04, 0.06]	1	0	0	0	0	0	0	0	1
(0.06, 0.08]	4	0	0	0	0	0	0	0	4
(0.08, 0.10]	5	2	5	0	0	0	0	0	12
(0.10, 0.12]	4	7	9	10	2	1	0	0	33
(0.12, 0.14]	0	4	7	5	1	1	0	0	18
(0.14, 0.16]	0	0	1	2	1	0	1	2	7
(0.16, 0.18]	0	0	0	1	0	1	0	0	2
(0.18, 0.24]	0	0	1	1	3	0	1	0	6
TOTAL	14	13	23	19	7	3	2	2	83

C) SUM07 (ppm-h)

HDM2 \ LOWER	0	10	20	30	40	50	60	70	TOTAL
(ppm) \ UPPER	10	20	30	40	50	60	70	130	
(0.04, 0.06]	1	0	0	0	0	0	0	0	1
(0.06, 0.08]	4	0	0	0	0	0	0	0	4
(0.08, 0.10]	8	3	1	0	0	0	0	0	12
(0.10, 0.12]	6	10	10	5	2	0	0	0	33
(0.12, 0.14]	1	8	6	1	1	1	0	0	18
(0.14, 0.16]	0	0	1	2	1	0	1	2	7
(0.16, 0.18]	0	0	1	0	0	0	1	0	2
(0.18, 0.24]	0	1	0	1	1	2	1	0	6
TOTAL	20	22	19	9	5	3	3	2	83

TABLE 4.5. (cont)

D) SUM06 (ppm-h)

M7 \ LOWER	0	14	28	42	56	70	84	98	TOTAL
(ppm) \ UPPER	14	28	42	56	70	84	98	168	
(0.024,0.030]	4	0	0	0	0	0	0	0	4
(0.030,0.036]	5	2	0	0	0	0	0	0	7
(0.036,0.042]	2	10	4	1	0	0	0	0	17
(0.042,0.048]	3	0	11	5	1	0	0	0	20
(0.048,0.054]	0	1	8	10	3	1	0	0	23
(0.054,0.060]	0	0	0	3	3	1	1	0	8
(0.060,0.066]	0	0	0	0	0	1	1	1	3
(0.066,0.078]	0	0	0	0	0	0	0	1	1
TOTAL	14	13	23	19	7	3	2	2	83

E) SUM07 (ppm-h)

M7 \ LOWER	0	10	20	30	40	50	60	70	TOTAL
(ppm) \ UPPER	10	20	30	40	50	60	70	130	
(0.024,0.030]	4	0	0	0	0	0	0	0	4
(0.030,0.036]	7	0	0	0	0	0	0	0	7
(0.036,0.042]	6	10	1	0	0	0	0	0	17
(0.042,0.048]	3	7	8	2	0	0	0	0	20
(0.048,0.054]	0	5	9	5	3	1	0	0	23
(0.054,0.060]	0	0	1	2	2	1	2	0	8
(0.060,0.066]	0	0	0	0	0	1	1	1	3
(0.066,0.078]	0	0	0	0	0	0	0	1	1
TOTAL	20	22	19	9	5	3	3	2	83

F) SUM06 (ppm-h)

SUM07 \ LOWER	0	14	28	42	56	70	84	98	TOTAL
(ppm-h) \UPPER	14	28	42	56	70	84	98	168	
(0, 10]	14	5	1	0	0	0	0	0	20
(10, 20]	0	8	14	0	0	0	0	0	22
(20, 30]	0	0	8	11	0	0	0	0	19
(30, 40]	0	0	0	8	1	0	0	0	9
(40, 50]	0	0	0	0	4	1	0	0	5
(50, 60]	0	0	0	0	2	1	0	0	3
(60, 70]	0	0	0	0	0	1	2	0	3
(70,130]	0	0	0	0	0	0	0	2	2
TOTAL	14	13	23	19	7	3	2	2	83

(6

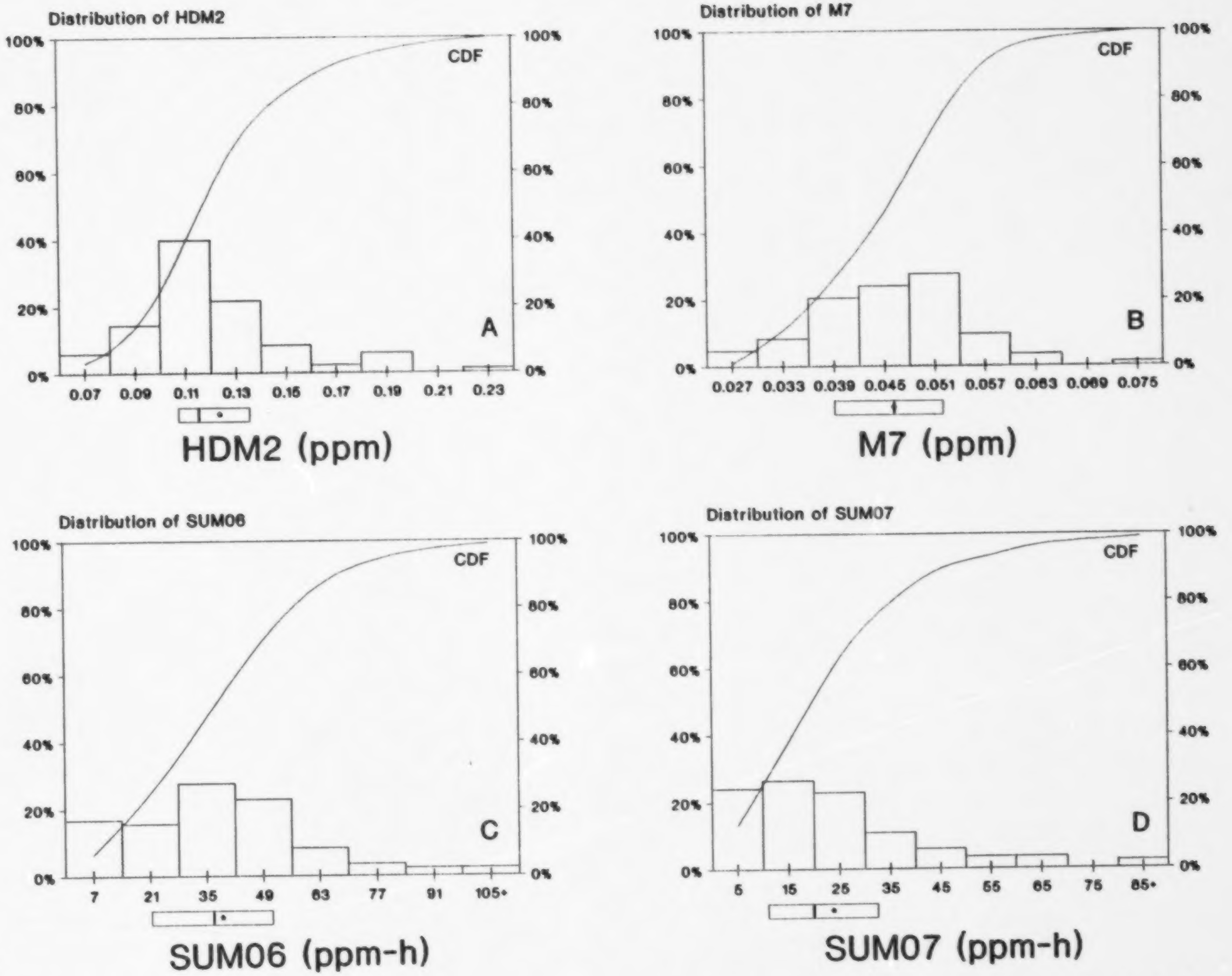


Figure 4.1. Cumulative distribution functions (CDFs) and frequency distributions of (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07 from 83 agricultural/ remote SAROAD site-years of ambient data. Except for HDM2, the exposure indices were calculated using hourly concentrations from May through September.

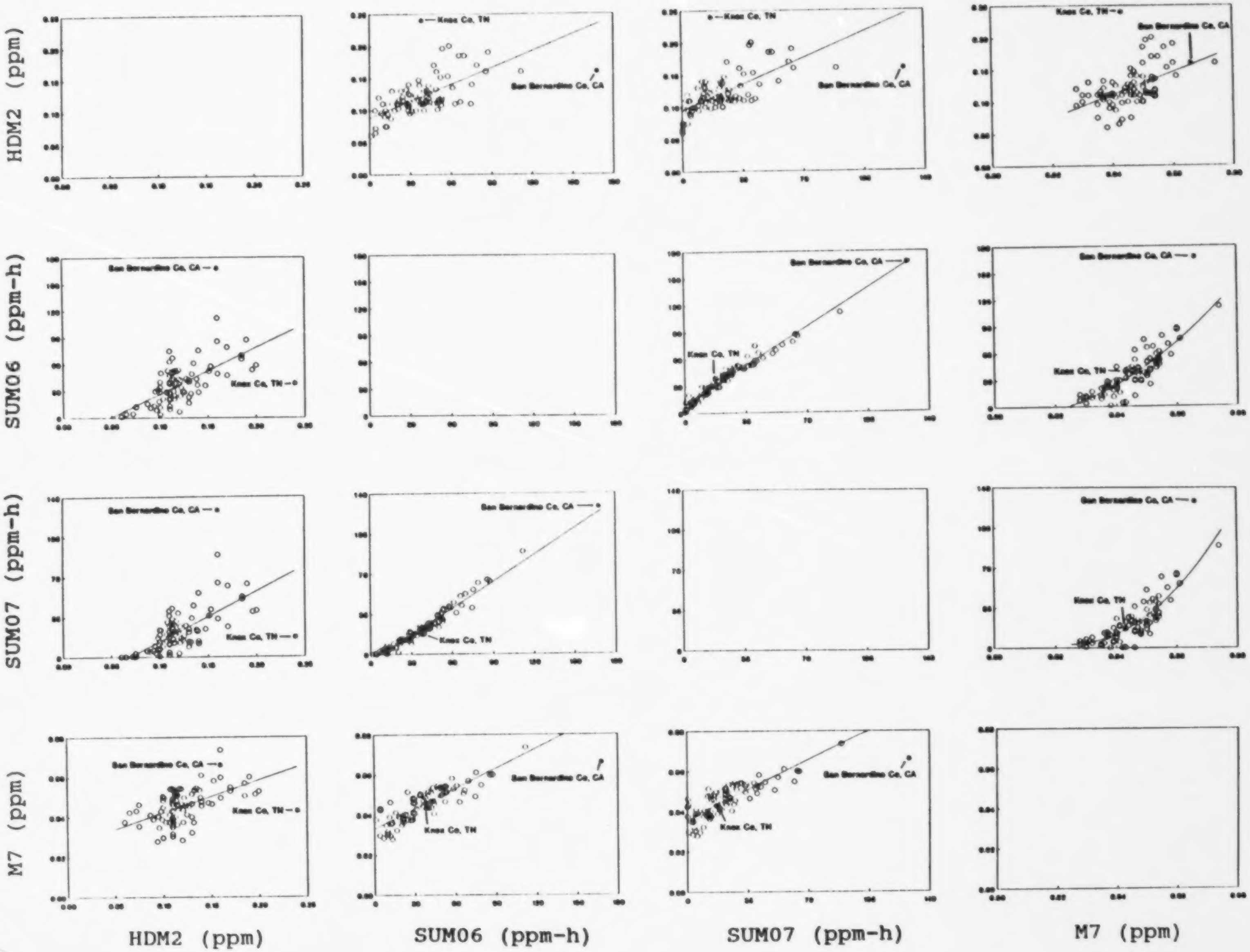


Figure 4.2. Scatterplot matrix of the four exposure indices (HDM2, SUM06, SUM07, and M7) from 83 agricultural/remote SAROAD site-years.

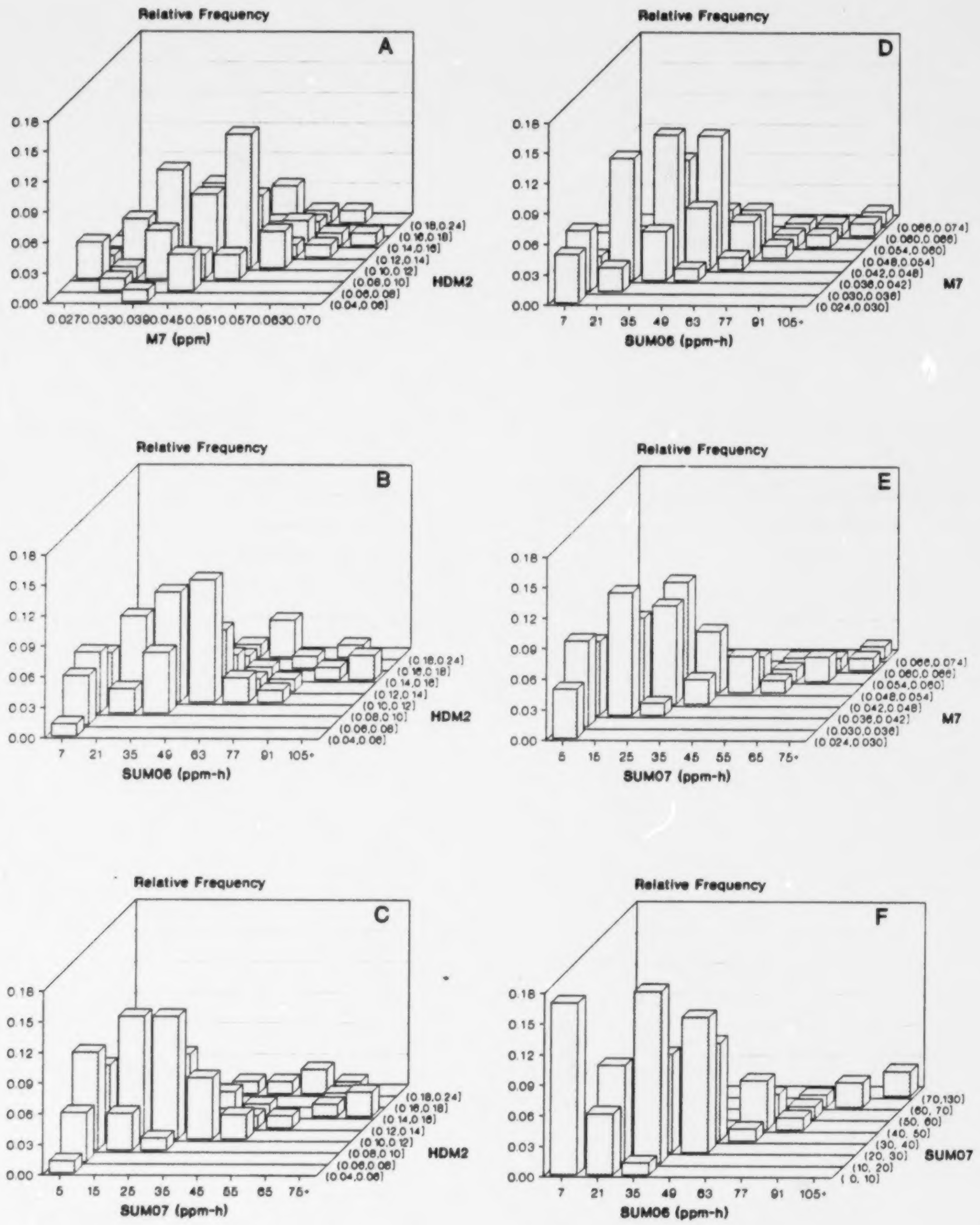


Figure 4.3. Joint distributions of co-occurrence of (A) M7 and HDM2, (B) SUM06 and HDM2, (C) SUM07 and HDM2, (D) SUM06 and M7, (E) SUM07 and M7, and (F) SUM06 and SUM07, for 83 agricultural/remote SAROAD site-years of data.

60

REFERENCES

Adams, R. M., S. A. Hamilton, and B. A. McCarl. (1985). An assessment of the economic effects of ozone on U.S. agriculture. J. Air Pollut. Control Assn., 35, 938-943.

Conover, W. J. (1971). Practical Nonparametric Statistics. New York, John Wiley & Sons, 461 pp.

Cure, W. W., J. S. Sanders, and A. S. Heagle. (1986). Crop yield response predicted with different characterizations of the same ozone treatments. J. Environ. Qual., 15, 251-254.

Heagle, A. S., and W. W. Heck. (1980). Field methods to assess crop losses due to oxidant air pollutants. In: Crop Loss Assessment: Proceedings of E.C. Stakman Commemorative Symposium, ed. by P.S. Teng and S.V. Krupa, Misc. Publication # 7, University of Minnesota, St. Paul, 296-305.

Heagle, A. S., W. W. Heck, J. O. Rawlings, and R. B. Philbeck. (1983). Effects of chronic doses of ozone and sulfur dioxide on injury and yield of soybeans in open-top field chambers. Crop Sci., 23, 1184-1191.

Heagle, A. S., V. M. Lesser, J. O. Rawlings, W. W. Heck, and R. B. Philbeck. (1986). Response of soybeans to chronic doses of ozone applied as constant or proportional additions to ambient air. Phytopathology, 76, 51-56.

Heck, W. W., O. C. Taylor, R. M. Adams, G. Bingham, J. Miller, E. Preston, and L. Weinstein. (1982). Assessment of crop loss from ozone. J. Air Pollut. Control Assn., 32, 353-361.

Heck, W. W., R. M. Adams, W. W. Cure, A. S. Heagle, H. E. Heggestad, R. J. Kohut, L. W. Kress, J. O. Rawlings, and O. C. Taylor. (1983). A reassessment of crop loss from ozone. Environ. Sci. Technol., 17, 573A-581A.

Heck, W. W., W. W. Cure, J. O. Rawlings, L. J. Zaragoza, A. S. Heagle, H. E. Heggestad, R. J. Kohut, L. W. Kress, and P. J. Temple. (1984a). Assessing impacts of ozone on agricultural crops: I. Overview. J. Air Pollut. Control Assn., 34, 729-735.

Heck, W. W., W. W. Cure, J. O. Rawlings, L. J. Zaragoza, A. S. Heagle, H. E. Heggestad, R. J. Kohut, L. W. Kress, and P. J. Temple. (1984b). Assessing impacts of ozone on agricultural crops: II. Crop yield functions and alternative exposure statistics. J. Air Pollut. Control Assn., 34, 810-817.

Hogsett, W. E., D. T. Tingey, and S. R. Holman. (1985). A pollutant exposure control system for determination of the effects of exposure regimes on plant growth. *Atmos. Environ.*, 19, 1135-1145.

Hogsett, W. E., D. T. Tingey, and E. H. Lee. (1988). Exposure indices: Concepts for development and evaluation of their use. In: *Assessment of Crop Loss from Air Pollutants*, ed. W. W. Heck, O. C. Taylor, and D. T. Tingey, New York, Elsevier Applied Science, 107-140.

Iman, R. L., and W. J. Conover. (1987). A measure of top-down correlation. *Technometrics*, 29, 351-357.

Johnson, T., A. Ferdo, and J. Capel. (1986). Relationships between selected three-month air quality indicators used to predict ozone damage to crops and woodland. Report to U.S. EPA Office of Air Quality Planning and Standards, Research Triangle Park, NC.

Kohut, R. J., R. G. Admundsen, J. A. Laurence, L. Colavito, P. van Leuken, and P. King. (1987). Effects of ozone and sulfur dioxide on yield of winter wheat. *Phytopathology*, 77, 71-74.

Kopp, R. J., W. J. Vaughan, and M. Hazilla. (1984). *Agricultural sector benefits analysis for ozone: Methods evaluation and demonstration*. Research Triangle Park, NC, U.S. Environmental Protection Agency. Available from NTIS, Springfield, VA; PB85-119477/XAB. EPA-450/5-84-003.

Kress, L. W., and J. E. Miller. (1983). Impact of ozone on soybean yield. *J. Environ. Qual.*, 12, 276-281.

Kress, L. W., and J. E. Miller. (1985a). Impact of ozone on field-corn yield. *Can. J. Bot.*, 63, 2408-2415.

Kress, L. W., and J. E. Miller. (1985b). Impact of ozone on grain sorghum yield. *Water, Air, Soil Pollut.*, 25, 377-390.

Kress, L. W., J. E. Miller, and H. J. Smith. (1985). Impact of ozone on winter wheat yield. *Environ. Exp. Bot.*, 25, 211-228.

Larsen, R. I., A. S. Heagle, and W. W. Heck. (1983). An air quality data analysis system for interrelating effects, standards, and needed source reductions: Part 8. An O_3-SO_2 leaf injury mathematical model. *J. Air Pollut. Control Assn.*, 33, 198-207.

Larsen, R. I., and W. W. Heck. (1984). An air quality data analysis system for interrelating effects, standards, and needed source reductions: Part 8. An effective mean O_3 crop reduction mathematical model. J. Air Pollut. Control Assn., 34, 1023-1034.

Larsen, R. I., and W. W. Heck. (1985). An air quality data analysis system for interrelating effects, standards, and needed source reductions: Part 9. Calculating effective ambient air quality parameters. J. Air Pollut. Control Assn., 35, 1274-1279.

Larsen, R. I., T. R. McCurdy, P. M. Johnson, and W. W. Heck. (1989). An air quality data analysis system for interrelating effects, standards, and needed source reductions: Part 10. Potential ambient O_3 standards to limit soybean crop reduction. J. Air Pollut. Control Assn., (in press).

Lee, E. H., D. T. Tingey, and W. E. Hogsett. (1987). Selection of the best exposure-response model using various 7-hour ozone exposure statistics. U.S. EPA, Office of Air Quality Planning and Standards, Research Triangle Park, NC.

Lee, E. H., D. T. Tingey and W. E. Hogsett. (1988a). Evaluation of ozone exposure indices for relating exposure to plant production and for estimating agricultural losses. U.S. EPA, Office of Air Quality Planning and Standards, Research Triangle Park, NC.

Lee, E. H., D. T. Tingey and W. E. Hogsett. (1988b). Evaluation of ozone exposure indices in exposure-response modeling. Environ. Pollut., 53, 43-62.

Lefohn, A. S., and D. T. Tingey. (1985). Comments on injury response of Phaseolus vulgaris to ozone flux density. Atmos. Environ., 21, 561-568.

Lefohn, A. S., J. A. Laurence, and R. J. Kohut. (1988). A comparison of indices that describe the relationship between exposure to ozone and reduction in the yield of agricultural crops. Atmos. Environ., 22, 1229-1240.

McCune, D. C. (1980). Terrestrial vegetation air pollutant interactions: Gaseous pollutant hydrogen flouride and sulfur dioxide. In: Proceedings of the International Conference on Air Pollutants and Their Effects on the Terrestrial Ecosystem, ed. S. V. Krupa and A. H. Legge, New York, Wiley-Interscience.

Musselman, R. C., R. J. Oshima, and R. E. Gallavan. (1983). Significance of pollutant concentration distribution in the response of 'red kidney' beans to ozone. *J. Am. Soc. Hortic. Sci.*, 108, 645-648.

Oshima, R. J. (1985). An assessment of experimental methods in research into air-pollutant effects on vegetation. In: *Sulfur Dioxide and Vegetation. Physiology, Ecology, and Policy Issues*, ed. by W. E. Winner, H. A. Mooney, and R. A. Goldstein, Stanford, Stanford University Press, 495-506.

Rawlings, J. O., and W. W. Cure. (1985). The Weibull function as a dose-response model to describe ozone effects on crop yields. *Crop Sci*., 25, 807-814.

Rawlings, J. O., V. M. Lesser, A. S. Heagle, and W. W. Heck. (1988). Alternative ozone dose metrics to characterize ozone impact on crop yield loss. *J. Environ. Qual.*, 17, 285-291.

SAS Institute, Inc. (1985). *SAS*, SAS Institute, Inc, Cary, NC.

Snedecor, G. W., and W. G. Cochran. (1980). *Statistical Methods 7th Edition*. Ames, Iowa, The Iowa State University Press, 507 pp.

Temple, P. J., O. C. Taylor, and L. F. Benoit. (1985). Cotton yield responses to ozone as mediated by soil moisture and evapotranspiration. *J. Environ. Qual.*, 14, 55-60.

Tingey, D. T. (1984). The effects of ozone on plants in the United States. In: *Proceedings of the International Workshop on the Evaluation of the Effects of Photochemical Oxidants on Human Health, Agricultural Crops, Forestry, Materials, and Visibility* ed. by P. Greenfelt, Goteborg, Sweden, Swedish Environmental Research Institute. 60-71.

Tingey, D. T., W. E. Hogsett, and E. H. Lee. (1988). Analysis of crop loss for alternative ozone exposure indices. Accepted for publication in *Proceedings of the 3rd U.S.-Dutch International Symposium on Atmospheric Ozone Research and its Policy Implications*. Elsevier Science Publishers.

U.S. Environmental Protection Agency. (1982). *Air quality data - 1981 annual statistics including summaries with reference to standards*. Research Triangle Park, NC, U.S. Environmental Protection Agency. EPA-450/4-82-007.

U.S. Environmental Protection Agency. (1986). *Air quality criteria for ozone and other photochemical oxidants, volume I*. Research Triangle Park, NC, U.S. Environmental Protection Agency. EPA-600/8-84-020aF.

U.S. Environmental Protection Agency. (1987). *Review of the national ambient air quality standards for ozone. Preliminary assessment of scientific and technical information*. Research Triangle Park, NC, U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards.

APPENDIX A

ABBREVIATIONS, CODE NAMES, AND FORMULAE

1) EXPOSURE INDICES

a) M7 = seasonal mean of 7-h daily means (0900-1600)

b) M1 = seasonal mean of 1-h daily means (0900-1600)

c) P1 = seasonal maximum of 1-h maximum daily means

d) P7 = seasonal maximum of 7-h maximum daily means

e) HDM2 = second highest 1-h maximum daily means

f) PERxx = the xx-th percentile of hourly distribution

g) TOTDOSE = seasonal sum of hourly concentrations
= $\Sigma O(i)$ where $O(i)$ is the hourly concentration

h) SUMxx = seasonal sum of hourly concentrations $\geq$ 0.xx ppm
= $\Sigma O(i) * I[O(i) \geq 0.xx]$

where $I[O(i) \geq 0.xx] = 1$ if $O(i) \geq 0.xx$ and $= 0$ otherwise, i.e., the indicator function for $O(i) \geq 0.xx$.

i) AOTxx = censored seasonal sum of hourly concentrations $\geq$ 0.xx ppm
= $\Sigma(O(i) - 0.xx) * I[O(i) \geq 0.xx]$

j) HRSxx = seasonal total number of hours $\geq$ 0.xx ppm
= $\Sigma I[O(i) \geq 0.xx]$

k) NUMEPxx = total number of episodes of 0.xx ppm or greater

l) AVGEPxx = average number of hours in an episode of concentrations of 0.xx or greater

m) DAYBETxx = average number of days between episodes of 0.xx ppm or greater

n) EFFMEAN = effective mean of Larsen and Heck (1984)
= $[\Sigma O(i)^{1/v}/N]^{v}$

where N=total number of hours and v=0.376

o) TIMPACT = total impact of Larsen *et al*. (1983)
= $\Sigma O(i)^{2.525}$

p) SIGMOID = sigmoid-weighted cumulative index

$$= \Sigma\{O(i) * W[O(i)]\}$$

where $W[O(i)]=[1 + M * \exp\{-c*O(i)\}]^{-1}$, i.e., the sigmoid function, SIG3, with c=100 and M=500.

q) ALLOMETRIC = allometric-weighted cumulative index

$$= \Sigma\{O(i) * W[O(i)]\}$$

where $W[O(i)]=O(i)^{(a-1)}$, i.e., the allometric function with a=3.5.

r) GPWCI = generalized phenologically weighted cumulative impact index

$$= \Sigma\{O(i) * W[O(i)] * PW(t) * F[O(i),O(i-1)]\}$$

where $W[O(i)]$ is a sigmoid or allometric function; $PW(t)$ is a gamma function scaled to a maximum value of one, i.e.,

$$PW(t) \propto (T0-t)^{(a-1)} * \exp[-b*(T0-t)]$$

where T0=final harvest day (or final harvest day + 10 for the shifted gamma function) and t=exposure day; and $F[O(i), O(i-1)]$ is the intra-episodic weighting function, i.e.,

$$\begin{aligned} F[O(i), O(i-1)] &= 0.8 \text{ if } O(i)\geq 0.xx \text{ ppm and } O(i-1)\geq 0.xx \text{ ppm} \\ &= 1.0 \text{ otherwise} \end{aligned}$$

2) WEIBULL EXPOSURE-RESPONSE MODEL

The Weibull exposure-response model relating plant response (Y) to an exposure index (X) and, possibly, a covariate (Z) is

$$Y = (A + D*Z) * \exp[-(X/B)^{C}] + e \qquad (1)$$

where e is the stochastic model error, and A, B, C, and D are unknown model parameters. The Weibull model for log transformed data is

$$\ln(Y) = A' + D'*Z + B'*X^{C'} + e \qquad (2)$$

where A', B', C', and D' are the unknown model parameters. When no covariate term is specified (D=0 in equation (1) and D'=0 in equation (2)), the model parameters in equation (2) are a reparametrization of model parameters in equation (1).

Namely,

$$A' = \ln(A),$$

$$B' = -(1/B)^{C},$$

and (3)

$$C' = C.$$

3) PREDICTED RELATIVE YIELD LOSS AND ESTIMATED EXPOSURE LEVELS

For the Weibull model in equation (1), the predicted relative yield loss (PRYL) calculated at an exposure level (X) is

$$\mathrm{PRYL}(X) = 1 - \exp[-(X/B)^{C}]. \qquad (4)$$

Conversely, for a fixed percent yield loss, say s%, the associated exposure level (X) from equation (4) is

$$X(s) = B*[-\ln(1-s/100)]^{(1/C)}. \qquad (5)$$

For log transformed plant response data and no covariate (Z), the model parameter estimates in equation (2) are reparametrized using equivalencies in equation (3). The predicted relative yield loss for given exposure level (X) is calculated using equation (4). The estimated exposure level associated with a yield loss of s% is calculated using equation (5).

APPENDIX B

AMBIENT AIR QUALITY INDICATORS AND ASSOCIATED PREDICTED RELATIVE YIELD LOSSES FOR AGRICULTURAL/REMOTE SITES IN THE U.S.

B.1. INTRODUCTION

Interrelating of experimental and ambient O_3 exposure data to identify appropriate air quality indicators necessitates consideration of the aggregation periods used to express pollutant exposure. The objective of this section is to interrelate the exposure-response functions derived in Chapter 2 that were used to estimate predicted relative yield losses (PRYLs) in Chapter 3 and ambient air quality data from 83 non-urban site-years (Chapter 4). The key issue is the resolution of the disparity between the aggregation periods used to express experimental and ambient pollutant concentrations.

Previous studies of ambient O_3 air quality data used different air quality indicators with various aggregation periods. Long-term averages were generally defined for a 1-month, 3-month, a growing season period (Johnson et al., 1986; Larsen et al., 1989; U.S. EPA, 1987). While peak indices were generally measured over the season (i.e., HDM2), the peaks most frequently occurred between May and September. The choice of aggregation periods has generally been a compromise between a short- and longer-term indicator.

Air quality data from field experiments generally report the O_3 exposures for a "season". The "season" varies according to species/cultivar, and with environmental and edaphic conditions, and covers the period when the plants receive experimental pollutant exposures (i.e., post-emergence to crop maturity). For example, in a series of field studies, exposure durations ranged from 36 to 111 days, with an average duration of 75 days (Table 3.1 in Chapter 3).

An aggregation period of five months (May to September) was recommended for describing ambient O_3 data (Chapter 4). The five-month period (1) captured 90% of the hourly concentrations of 0.06 ppm or higher, (2) included the EPA-defined O_3 season for most states, and (3) encompassed the majority of occurrences of the second highest daily 1-h maximum concentration (i.e., the current NAAQS). This time period also coincided with the agricultural season for the majority of annual crops grown in the United States.

The integrated exposure indices (SUM06 and SUM07) used to characterize experimental exposures are functions of exposure duration and concentration. Correspondingly, the exposure levels associated with various yield losses calculated from experimental data (Chapter 3) are functionally related to the exposure "seasons". An implicit assumption is that exposures of varying duration, with equal values of SUM06 (or SUM07) cause the same biological response (for the same crop/cultivar, environmental, and edaphic conditions). Experiments replicated in time and/or space, differing in exposure duration, should produce identical predicted relative yield losses using SUM06 or SUM07. Evidence to support this is given in Chapter 3. In replicated studies, wheat (VONA 82 and 83) and soybean (DAVIS 81 and 82) showed dissimilar responses when noncumulative indices (HDM2 and M7) were used (Figures 3.2 and 3.3). However, when integrated indices (SUM06 and SUM07) were used, predicted relative yield losses were similar (Figures 3.4 and 3.5). Differences in plant yield among similar experiments differing in exposure duration were largely explained by the dissimilarities in the integrated indices, i.e., exposure duration (Tingey et al., 1988).

The results indicated that the integrated indices, SUM06 and SUM07, captured the key components of the exposure (i.e., exposure duration and peak concentrations) and were adequate descriptors of plant response. The discrepancy between exposure "seasons" and between experimental and ambient aggregation periods were accounted for in the integrated indices. The discrepancy between experimental and ambient aggregation periods is not important for comparison of experimental exposure levels and ambient air quality indicators using noncumulative indices (e.g., HDM2, M7) since exposure duration is excluded from the calculations.

Crop yield losses were calculated, based on 83 site-years of ambient air monitoring data using the four exposure indices, the current secondary NAAQS (HDM2), the 7-h seasonal mean (M7), and the integrated indices (SUM06 and SUM07).

B.2. METHODS

Ambient air quality data for 83 agricultural/remote site-years were characterized using four exposure indices, HDM2, M7, SUM06, and SUM07 (see Table 4.1 in Chapter 4). For each site-year, crop yield losses for the 16 NCLAN studies were estimated using the Weibull exposure-response models (in Table 3.2 in Chapter 3) and the ambient air quality indicators. The results are presented in a series of graphs (Figures B.1-B.20).

74

B.3. RESULTS

The relative crop yield loss curves for each of the 16 NCLAN cases and 4 exposure indices are illustrated in Figures B.1-B.16. Superimposed on these graphs are the cumulative distribution functions (CDF) of exposure indices at the 83 site-years. The ambient values of SUM06 and SUM07 for a number of site-years exceeded the experimental range for these indices, requiring extrapolation of the crop yield loss curve (Table B.1). This portion of the curve is indicated by a dashed line.

To illustrate the use of these figures, consider the crop loss model fit using HDM2 for the 1980 soybean study at Argonne (Figure B.1A). Reading off the percent yield loss curve, a 20% yield loss is predicted to occur at approximately 0.12 ppm for HDM2. In turn, reference to the cumulative distribution function indicates that 60% of the site-years have HDM2 less than 0.12 ppm. That is, 60% of site-years meet the level of the current standard and are expected to have yield losses of 20% or less (based on the 1980 soybean study). These figures interrelate the exposure-response curves (discussed in Chapter 3) with the distributions of ambient air quality indicators at 83 non-urban site-years (Figure 4.1 in Chapter 4). The magnitude of estimated biological responses at the 83 site-years is presented in Table B.1.

For each of the 16 cases, the standard deviation of PRYLs at the 83 site-years is greater using the integrated indices (SUM06 and SUM07) than using the peak (HDM2) or mean (M7) indices (Table B.1). Considerable differences in median PRYLs at the 83 site-years exist among the peak, mean, and integrated indices. This is particularly evident in the cases with a "short season", i.e., A80SO (56 days), A82WH (56 days), A83WH (53 days), and I83WH (36 days).

To illustrate the range of estimated crop yield losses at the 83 areas, the 50th, 75th, and 85th percentiles of PRYL for 16 cases were calculated at each area (Figures B.17-B.19). The current standard, HDM2, was used to sort in ascending order the 83 site-years along the vertical axis. In this manner, a comparison was made of the estimated yield losses for sites with HDM2≤0.12 ppm and those sites with HDM2>0.12 ppm. Recall, in Chapter 3, that sites meeting or exceeding the current standard have significantly different values for M7, SUM06, and SUM07.

Figure B.17A shows that while the average M7 values for sites with HDM2≤0.12 ppm differ significantly from the M7 values for sites with HDM2>0.12 ppm (0.043 vs. 0.050 ppm), the median PRYLs do not differ. The Weibull PRYL curves using M7 (e.g., Figure B.1B) are fairly flat between 0.035 and 0.055 ppm (which captures 80% of the site-years; see Table 4.3 in Chapter 4). Based on M7, the majority of sites have estimated median yield losses around 10%.

Figure B.17B shows a wide range of responses using SUM06 between sites with HDM2≤0.12 ppm and those sites with HDM2>0.12 ppm. While sites with extremely high (low) values of HDM2 are associated with high (low) yield losses, sites with HDM2 values between 0.107 and 0.166 ppm (65% of the sites; see Table 4.3 in Chapter 4) are highly variable in SUM06 and in PRYLs. More importantly, sites meeting or exceeding the current O_3 standard (with HDM2 between 0.107 and 0.166 ppm) have similar responses.

Figure B.17C shows the estimated yield losses using SUM07 are smaller and less variable than the yield losses using SUM06. The problem of extrapolation is less pronounced using SUM07 than using SUM06 since SUM07 values are typically 20% less than SUM06 values. Otherwise, the pattern of responses is similar using SUM06 and SUM07. Sites with extremely high values of HDM2 have high yield losses. For sites with HDM2 between 0.107 and 0.166 ppm, the estimated yield losses do not differ significantly.

The more extreme crop yield losses at the 83 non-urban sites, described by the 75th and 85th percentiles of PRYLs for the 16 NCLAN cases (Figures B.18 and B.19, respectively), show similar patterns of response to the median PRYLs. Figures B.18 and B.19 indicate that the more sensitive crops have similar estimated responses at sites meeting or exceeding the level of the current O_3 standard (with HDM2 between 0.107 and 0.166 ppm). The 75th and 85th percentiles of PRYLs are more variable using SUM06 than using either SUM07 or M7. Sites with extremely high HDM2 values have 75th and 85th percentiles of PRYLs greater than 90% using SUM06 or SUM07. At these sites, crop yield losses of 90% or greater can be expected for 25% of all crops.

The range of crop yield losses is given by the 50th, 75th, and 85th percentiles of PRYLs in Figures B.20A-D for the four exposure indices, HDM2, M7, SUM06, and SUM07, respectively. The percentiles of the 4 air quality indicators at the 83 non-urban sites are also given and interrelate the empirical yield loss curves and the ambient air quality data. For example, an HDM2 value of 0.12 ppm results in a 19% or greater yield loss for half of the crops and roughly half of the 83 non-urban sites have a HDM2 value of 0.12 ppm or greater (Figure B.20A). An HDM2 value of 0.12 ppm results in a 40% or greater yield loss for 15% of the crops.

A median M7 value of 0.046 ppm corresponds to estimated losses of 6% or greater for 50% of all crops and estimated losses of 15% or greater for 15% of all crops (Figure B.20B). A median SUM06 value of 37 ppm-h corresponds to estimated yield losses of 25% or greater for 50% of all crops and estimated losses of 57% or greater for 15% of all crops (Figure B.20C). A median SUM07 value of 20 ppm-h corresponds to estimated yield losses of 12% or greater for 50% of all crops and estimated losses of 34% or greater for 15% of all crops (Figure B.20D).

B.4. DISCUSSION

Based on the median response of the 16 cases (including both O_3 sensitive and tolerant species/cultivars), a SUM06 (SUM07) exposure level of 21 (18) ppm-h is predicted to cause 10% yield loss (Chapter 3). Whether the exposure level is reached in 100 or 150 days is not important since the estimated yield losses are the same. The premise is that exposure duration is fully captured in the integrated indices (SUM06 and SUM07) and plant response is adequately explained by the integrated indices. For some crops, this level may be overly protective. Because the analysis is based on the median response, it will be under-protective for more sensitive crops.

The intent of a secondary O_3 NAAQS to protect vegetation may be interpreted as requiring protection of the most sensitive crops not just the median response (Oshima, 1985). In this case, the maximum crop loss should be used in deriving the exposure level of response. A more pragmatic view of the secondary standard would consider the distribution of effects and the proportion of areas affected (McCune, 1980). A probabilistic approach is needed to set an O_3 standard that protects the more sensitive crops against pollutant exposure with high probability. For example, based on the median relative yield losses, exposures with SUM06 of 21 ppm-h will result in yield losses of 10% (or less) for half of the crops. A more protective strategy might be based on the 75th percentile. Based on the 75th percentile, exposures with SUM06 of 17 ppm-h will result in yield losses of 10% (or less) for 75% of the crops (Figures B.20C-D). This will affect 80% of the non-urban sites using SUM06.

Based on the 75th percentile, exposures with HDM2 of 0.04 ppm will result in yield losses of 10% (or less) for 75% of the crops (Figures B.20A-B). This will affect all of the non-urban sites using HDM2. The exposure levels and proportion of sites affected using the 50th, 75th, and 85th percentile and mean for 10% PRYLs are summarized in Table B.2.

A problem in interrelating the ambient air quality indices with the crop yield loss functions exists. This is most evident in the SUM06 index, which has a number of ambient values outside the range of experimental O_3 exposures. On the other hand, the range of ambient values of SUM07 coincides fairly well with the experimental ranges. Extrapolation of relative yield losses using SUM07 is not a problem and does not affect the results significantly.

B.5. CONCLUSION

While the choice of aggregation period is important in capturing the key components of O_3 exposure, the aggregation period issue does not affect the interpretation of the exposure levels of specified yield losses for setting a secondary standard that protects vegetation.

Interrelating the crop loss functions and the ambient air quality indicators shows a range of possible plant responses at the 83 non-urban site-years, depending upon the exposure index. Median crop losses at the 83 site-years range from 10% using M7 to 27% using SUM06. The median and variation in crop loss should be considered in deriving appropriate ambient air quality indicators that protect vegetation, particularly crops that are sensitive to ozone. One means of achieving this goal is to set the secondary standard off the 75th percentile rather than the median predicted relative yield loss curve.

(71)

TABLE B.1.

Predicted relative yield losses (PRYL) for the 83 rural and agricultural sites based on the Weibull exposure-response models for 16 NCLAN studies.

		Maximum exposure levels from experimental ozone exposures				PRYL for M7			PRYL for HDM2			PRYL for SUM06			PRYL for SUM07		
ID	SPECIES	HDM2 (ppm)	M7 (ppm)	SUM06 (ppm-h)	SUM07 (ppm-h)	Range (%)	Median (%)	SD (%)	Range (%)	Median (%)	SD (%)	Range (%)	Median (%)	SD (%)	Range (%)	Median (%)	SD (%)
A80SO	CORSOY	0.200	0.115	42.9	42.5	4.3-26.4	11.4	4.0	3.8-70.7	17.9	12.0	0.2-99.5	40.9	24.2	0.0-91.9	22.5	19.5
A83CV	AMSOY	0.184	0.094	52.3	49.8	0.6-17.2	3.5	2.6	9.7-50.5	22.3	7.2	0.0-100.	16.9	23.3	0.0-88.3	18.8	15.9
A83CV	CORSOY	0.184	0.094	52.3	49.8	0.0- 9.9	0.6	1.3	4.5-58.5	16.7	9.5	0.0-100.	8.4	30.9	0.0-100.	1.8	21.0
R81SO	DAVIS	0.219	0.125	94.5	93.8	8.5-30.0	16.8	4.1	23.8-66.3	40.6	7.6	0.3-73.7	19.2	13.0	0.0-58.7	12.7	10.7
R82SO	DAVIS	0.193	0.111	67.3	65.5	2.9-25.5	9.4	4.0	8.5-57.7	22.8	8.8	0.1-94.4	25.0	18.9	0.0-79.4	14.7	15.0
A82WH	ABE	0.200	0.122	47.3	46.9	1.6-17.6	6.0	2.8	3.3-66.0	15.7	11.1	0.0-100.	31.0	26.8	0.0-96.7	13.3	20.0
A82WH	ARTHUR	0.200	0.122	47.3	46.9	2.0-17.9	6.5	2.8	3.2-64.0	15.1	10.7	0.0-100.	29.6	26.4	0.0-96.7	12.1	19.6
A83WH	ABE	0.175	0.095	33.5	32.7	0.0- 1.3	0.0	0.1	0.0-100.	0.8	14.1	0.0-100.	29.2	42.9	0.0-100.	1.0	29.9
A83WH	ARTHUR	0.175	0.095	33.5	32.7	0.2- 7.8	1.3	1.2	0.0-71.8	1.6	10.2	0.0-99.6	21.3	22.3	0.0-85.7	8.2	15.0
I82WH	VONA	0.181	0.096	41.7	40.8	44.4-78.8	62.3	7.0	54.1-95.6	77.5	7.6	3.9-99.8	74.1	22.7	0.0-99.2	51.9	25.7
I83WH	VONA	0.205	0.098	22.2	21.3	5.4-33.4	14.5	5.2	17.9-68.3	36.3	9.1	1.2-100.	73.6	26.9	0.0-99.3	46.2	26.0
A82SG	DEKALB	0.208	0.129	76.1	75.9	0.5- 4.4	1.6	0.7	0.4-21.5	2.7	3.3	0.0-59.0	2.8	7.8	0.0-33.8	1.0	4.5
A81MA	PAG-397	0.245	0.156	89.9	89.9	0.0- 3.5	0.5	0.5	0.0-55.4	1.3	7.6	0.0-100.	2.0	17.5	0.0-100.	2.0	17.5
A81MA	PIONEER	0.245	0.156	89.9	89.9	0.0-11.3	3.2	1.8	0.5-61.1	9.0	9.7	0.0-99.6	7.9	18.3	0.0-99.2	9.0	18.2
C81CO		0.247	0.137	67.5	65.7	5.2-26.0	12.4	3.9	20.1-61.2	35.6	7.4	0.0-98.5	24.2	21.2	0.0-97.9	26.8	21.0
C82CO		0.203	0.111	71.9	70.3	5.0-34.0	14.3	5.3	7.9-74.4	26.5	12.2	0.2-95.8	29.1	20.0	0.5-93.3	32.7	19.2
Average median and std. dev. (SD) PRYLs for 16 cases							10.3	3.0		21.4	9.3		27.2	22.7		17.2	18.7

Maximum ambient air quality indicators for 83 non-urban sites are 0.240 ppm, 167.9 ppm-h, and 128.2 ppm-h for HDM2, SUM06, and SUM07, respectively, exceeding the maximum ozone exposures used in the majority of NCLAN studies. This is a result of using an ambient averaging period from May to September (153 days) versus the average NCLAN season of 75 days (see Table 3.1).

TABLE B.2.

Interrelation of experimental and ambient air quality data to estimate the exposure levels associated with mean, 50th, 75th, and 85th percentiles of PRYLs of 10% and the proportion of sites affected.

	50th perc. PRYL		Mean PRYL		75th perc. PRYL		85th perc. PRYL	
	Exposure level	% sites affected	Exposure level	% sites affected	Exposure level	% sites affected	Exposure level	% sites affected
HDM2 (ppm)	0.089	91%	0.061	99%	0.041	100%	0.026	100%
M7 (ppm)	0.057	7%	0.045	56%	0.039	74%	0.035	89%
SUM06 (ppm-h)	21	77%	17	80%	17	80%	8	93%
SUM07 (ppm-h)	18	54%	14	62%	13	67%	6	80%

180

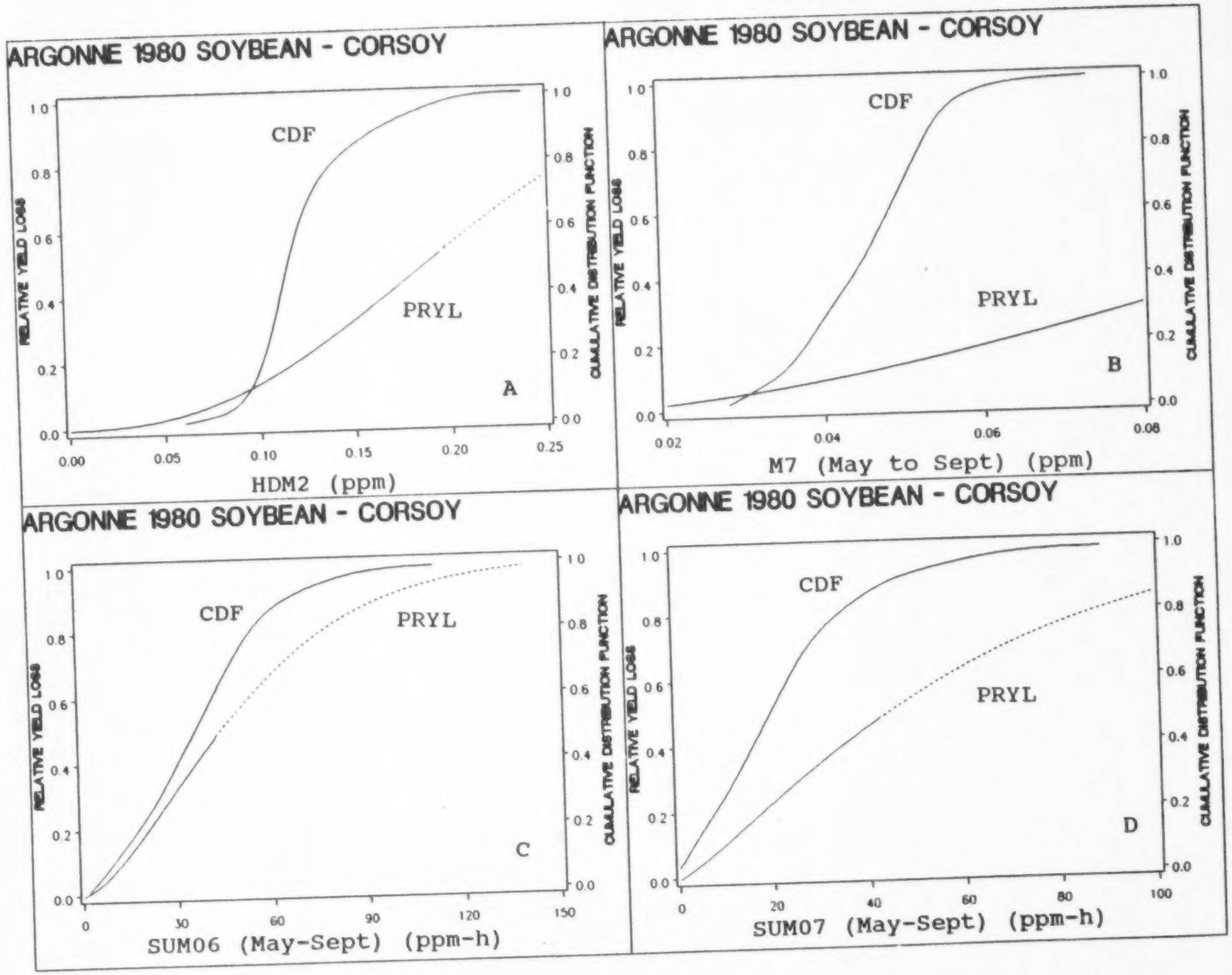


Figure B.1. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1980 soybean study at Argonne, IL using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

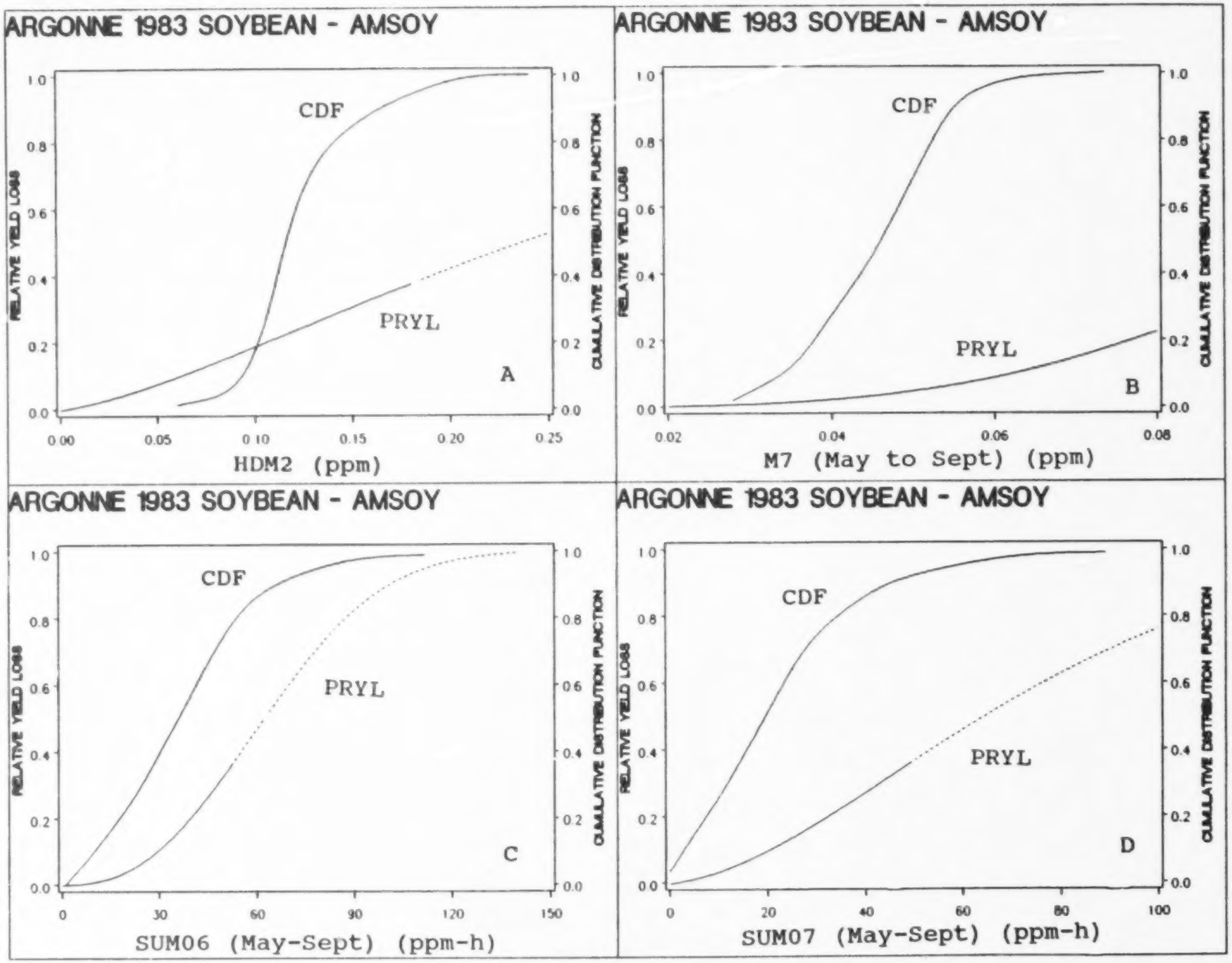


Figure B.2. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1983 soybean study (Amsoy) at Argonne, IL using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

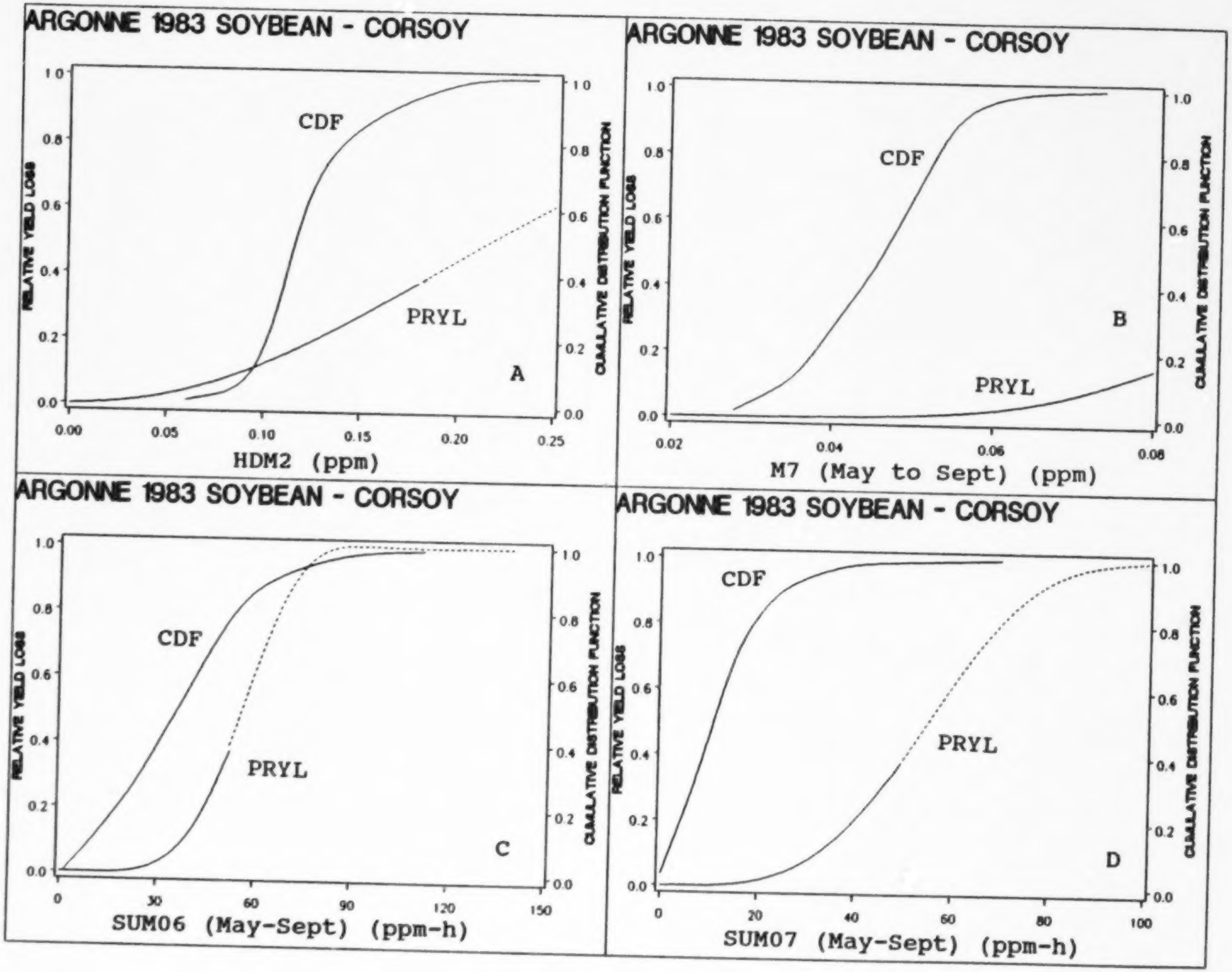


Figure B.3. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1983 soybean study (Corsoy) at Argonne, IL using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

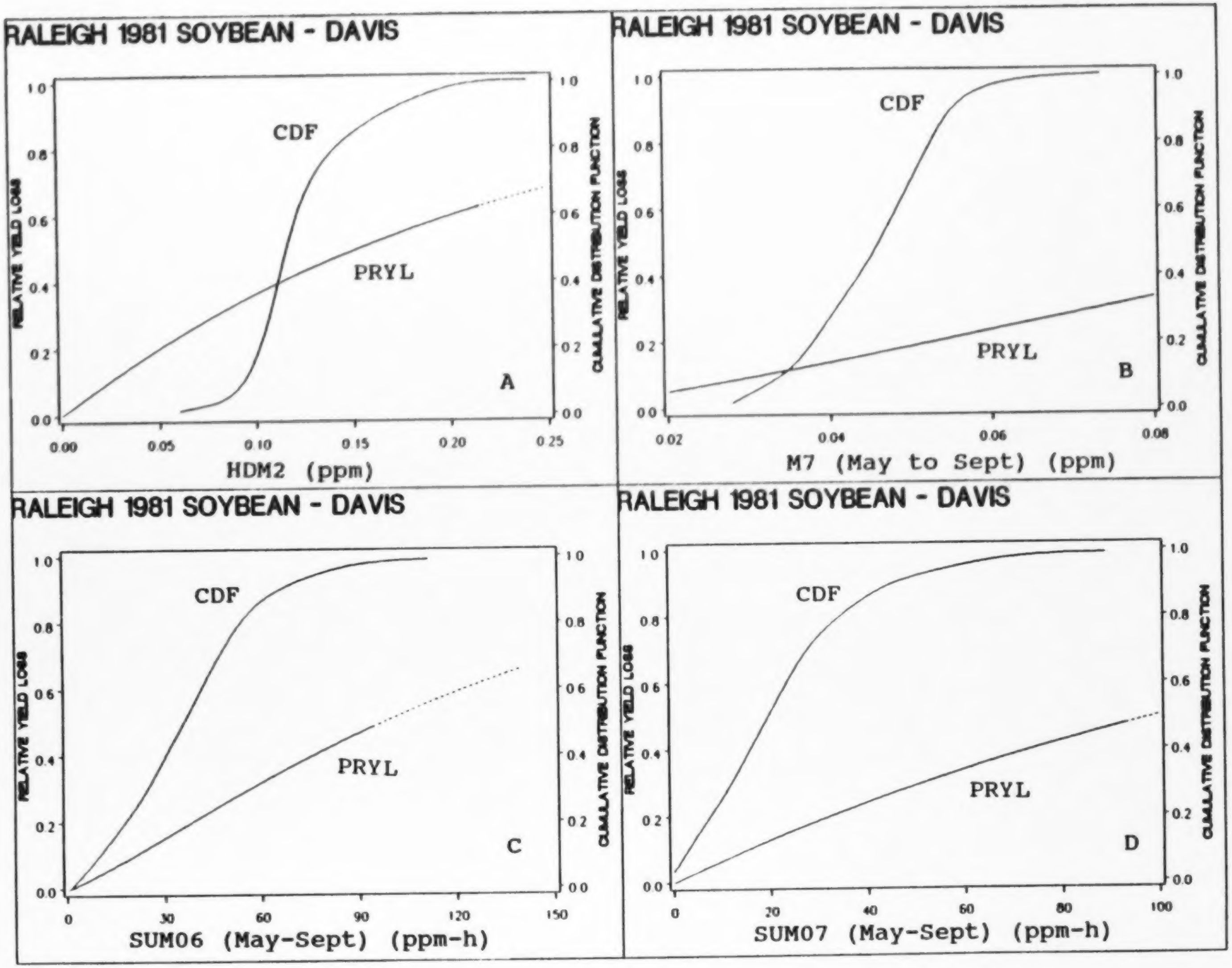


Figure B.4. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1981 soybean study at Raleigh, NC using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

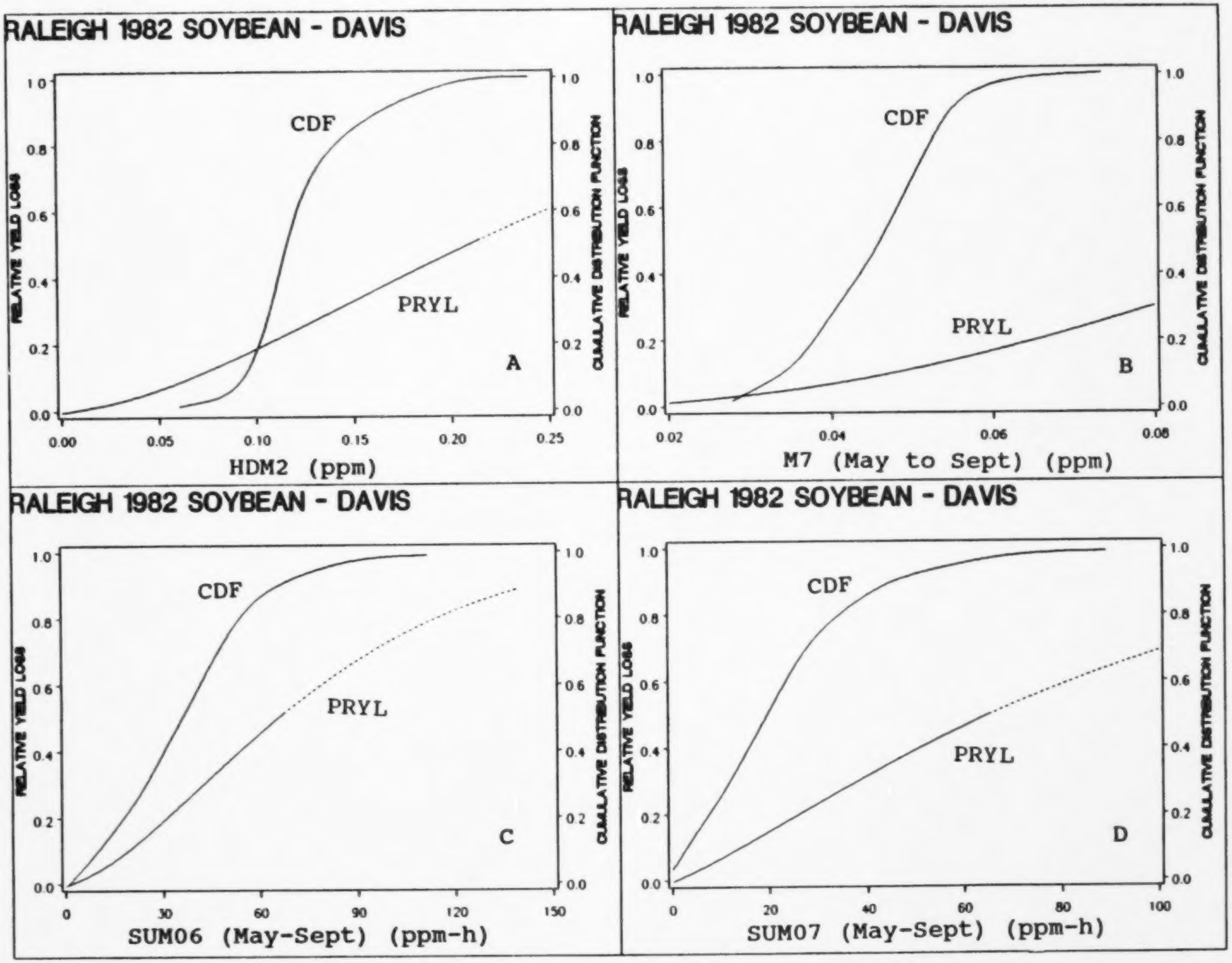


Figure B.5. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1982 soybean study at Raleigh, NC using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

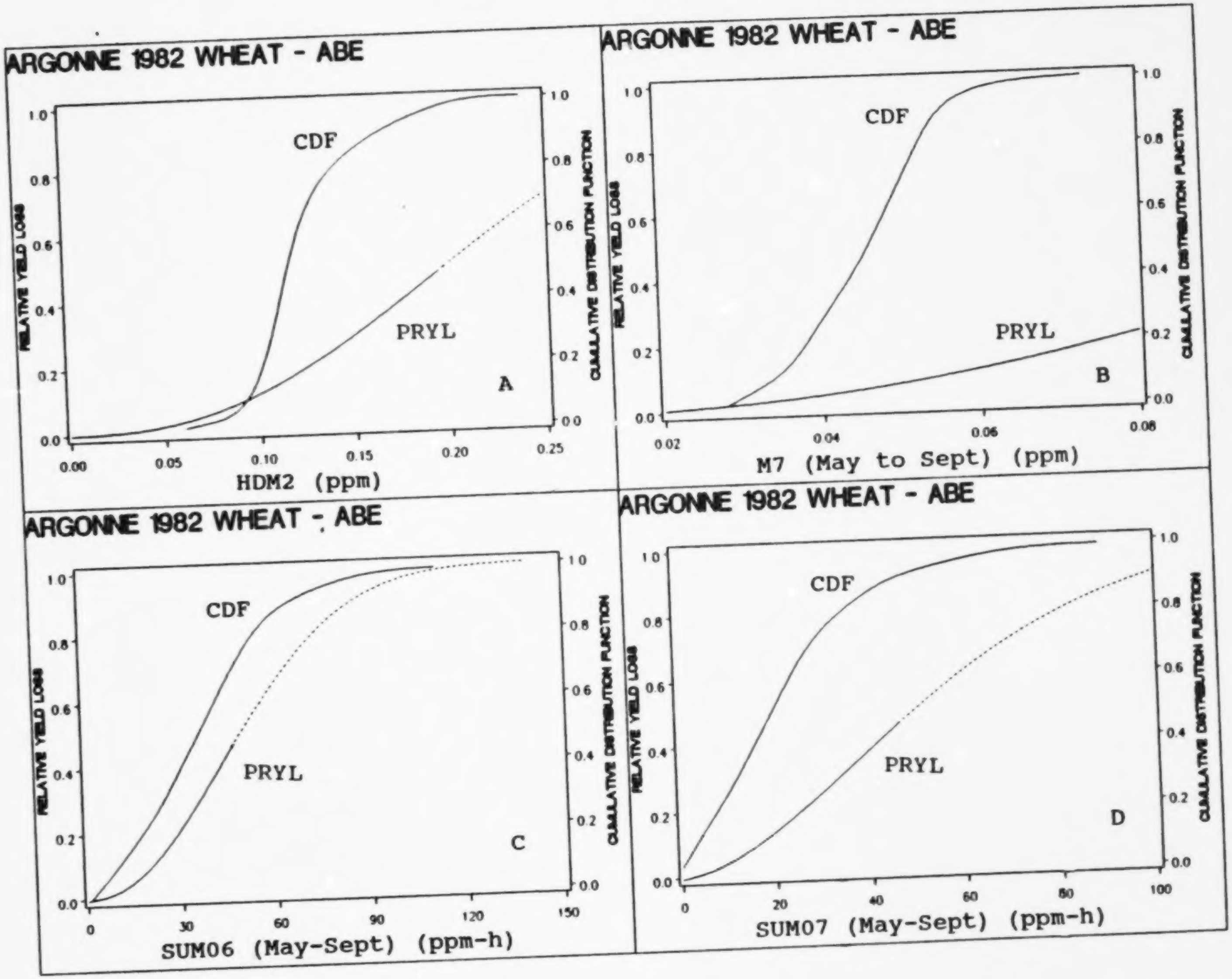


Figure B.6. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1982 wheat study (Abe) at Argonne, IL using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

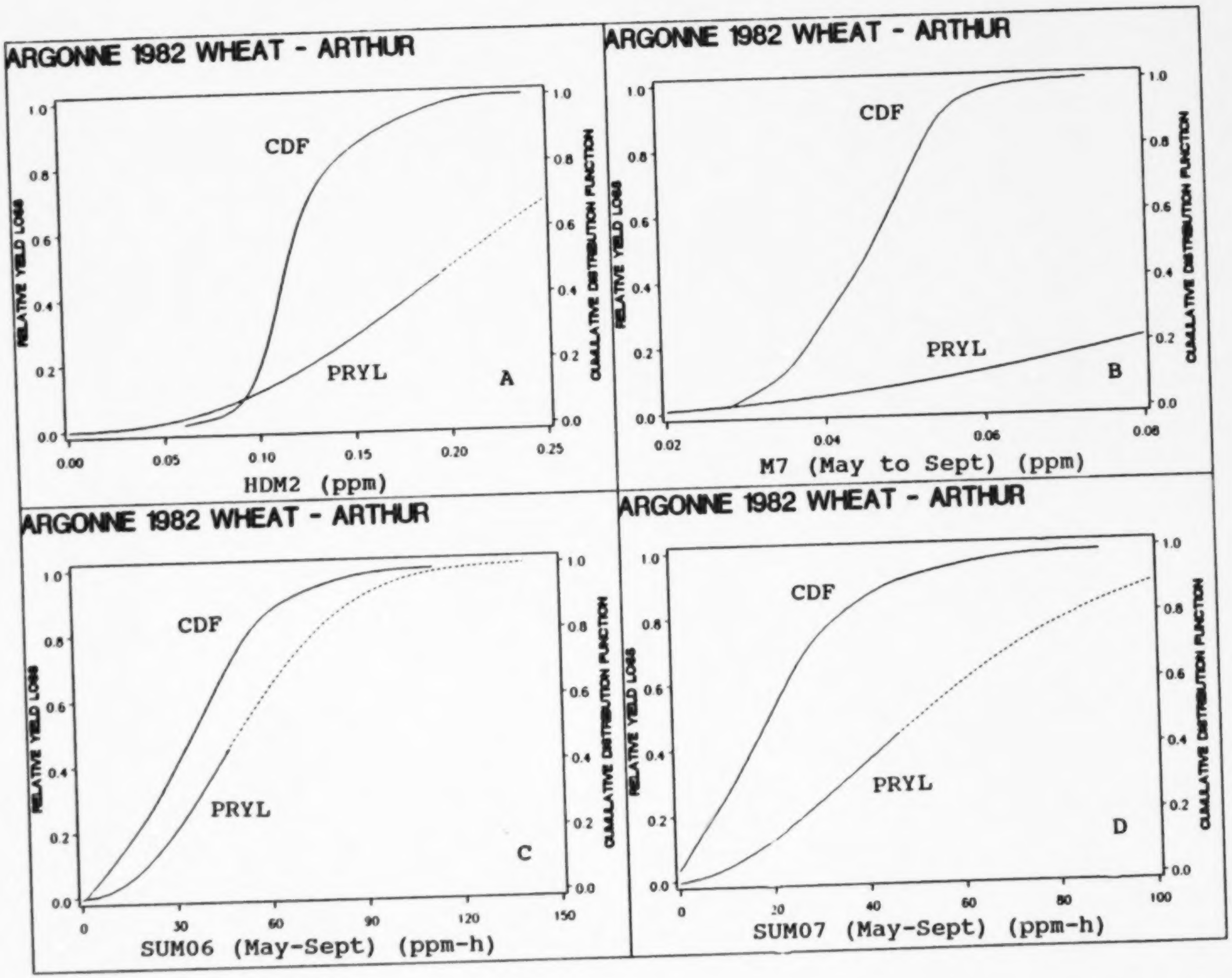
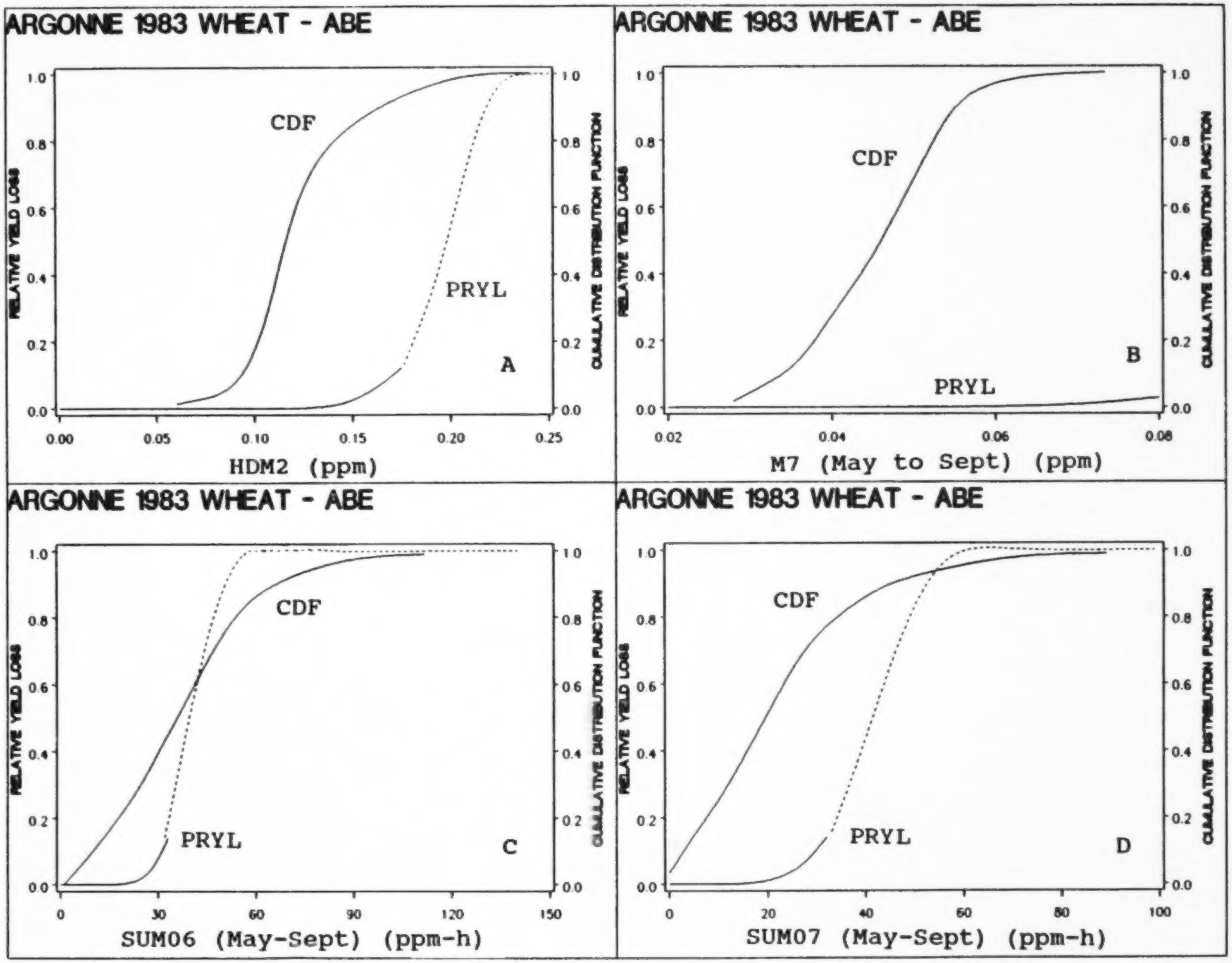


Figure B.7. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1982 wheat study (Arthur) at Argonne, IL using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.



ARGONNE 1983 WHEAT - ABE

ARGONNE 1983 WHEAT - ABE

ARGONNE 1983 WHEAT - ABE

ARGONNE 1983 WHEAT - ABE

Figure B.8. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1983 wheat study (Abe) at Argonne, IL using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

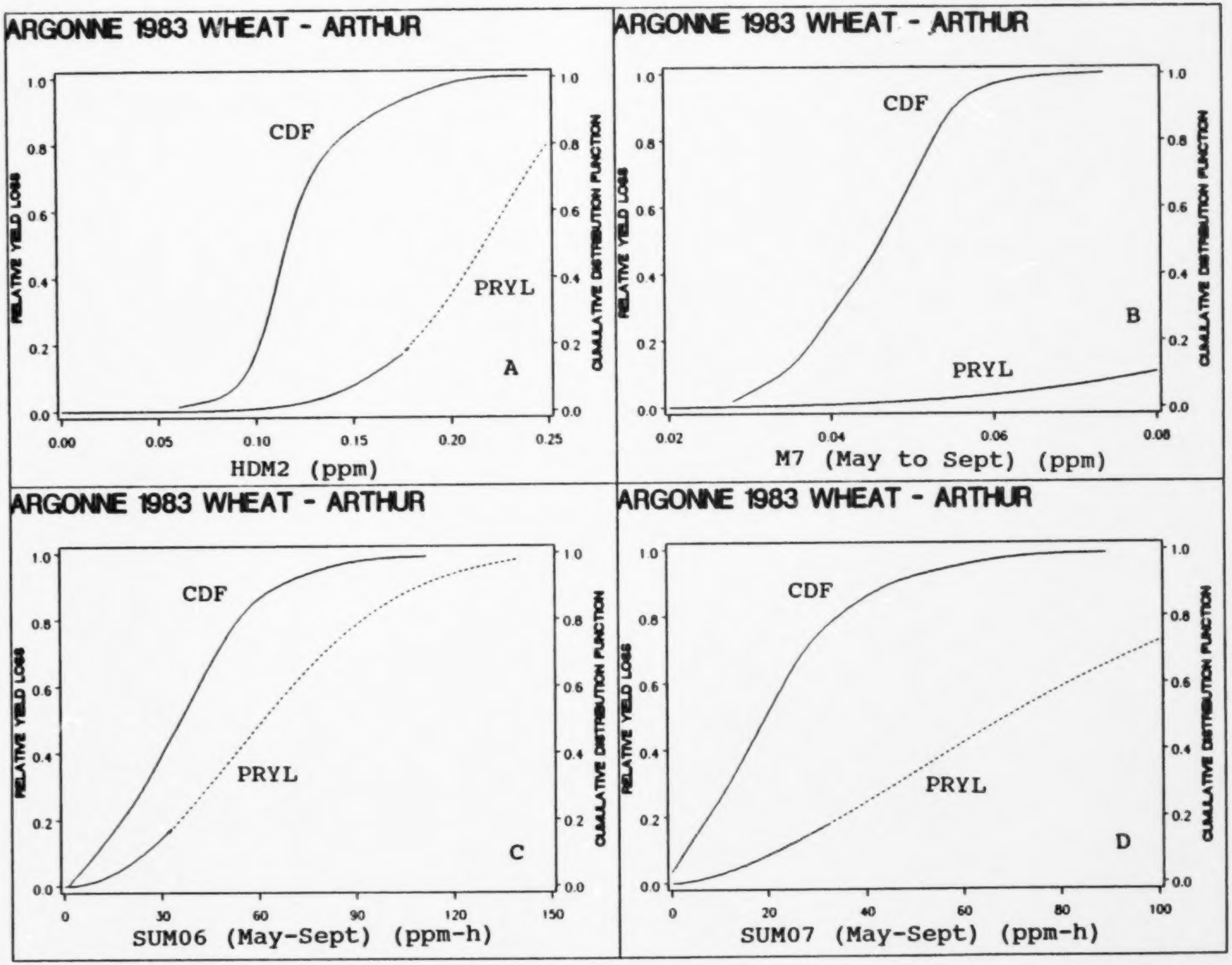


Figure B.9. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1983 wheat study (Arthur) at Argonne, IL using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

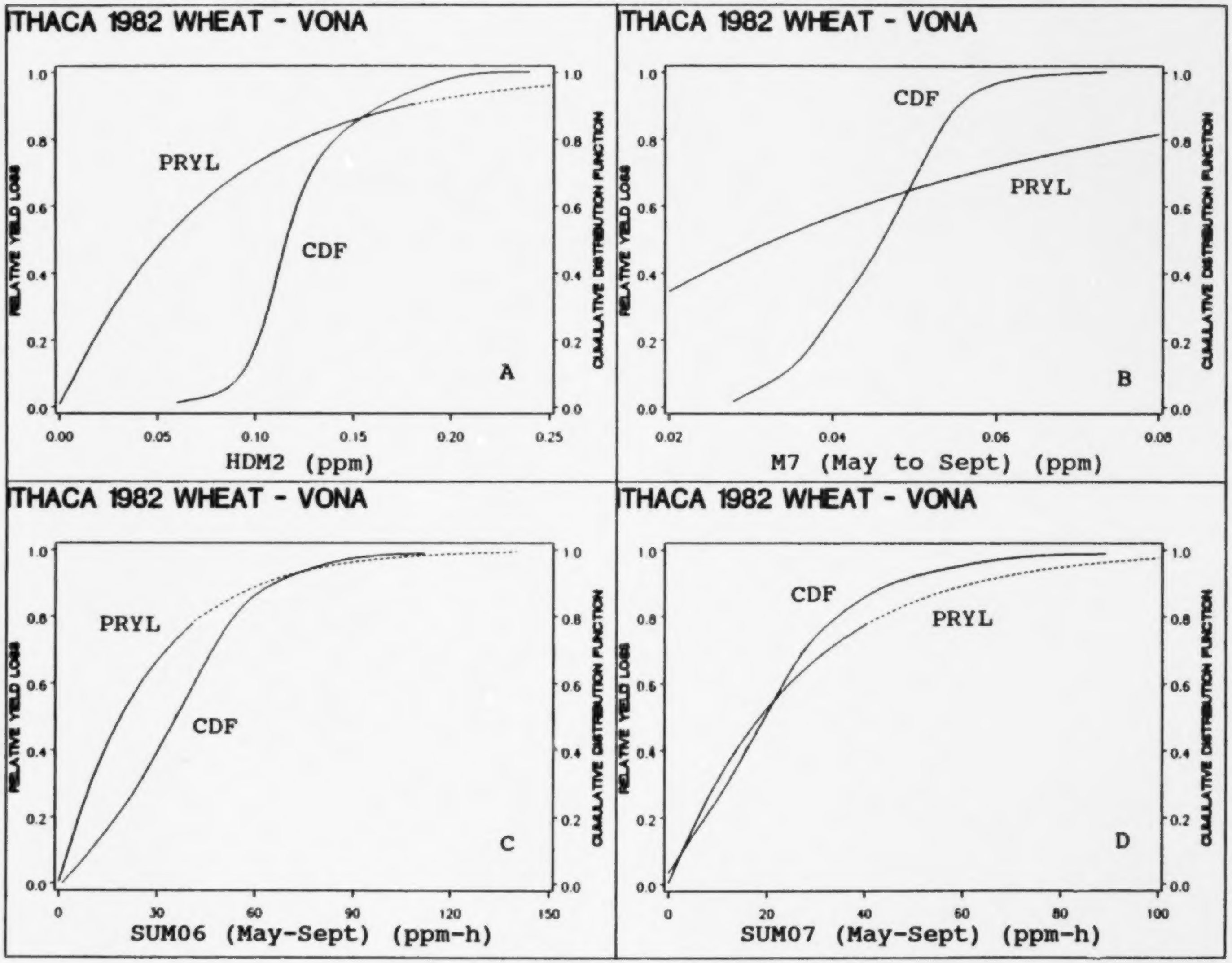


Figure B.10. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1982 wheat study at Ithaca, NY using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

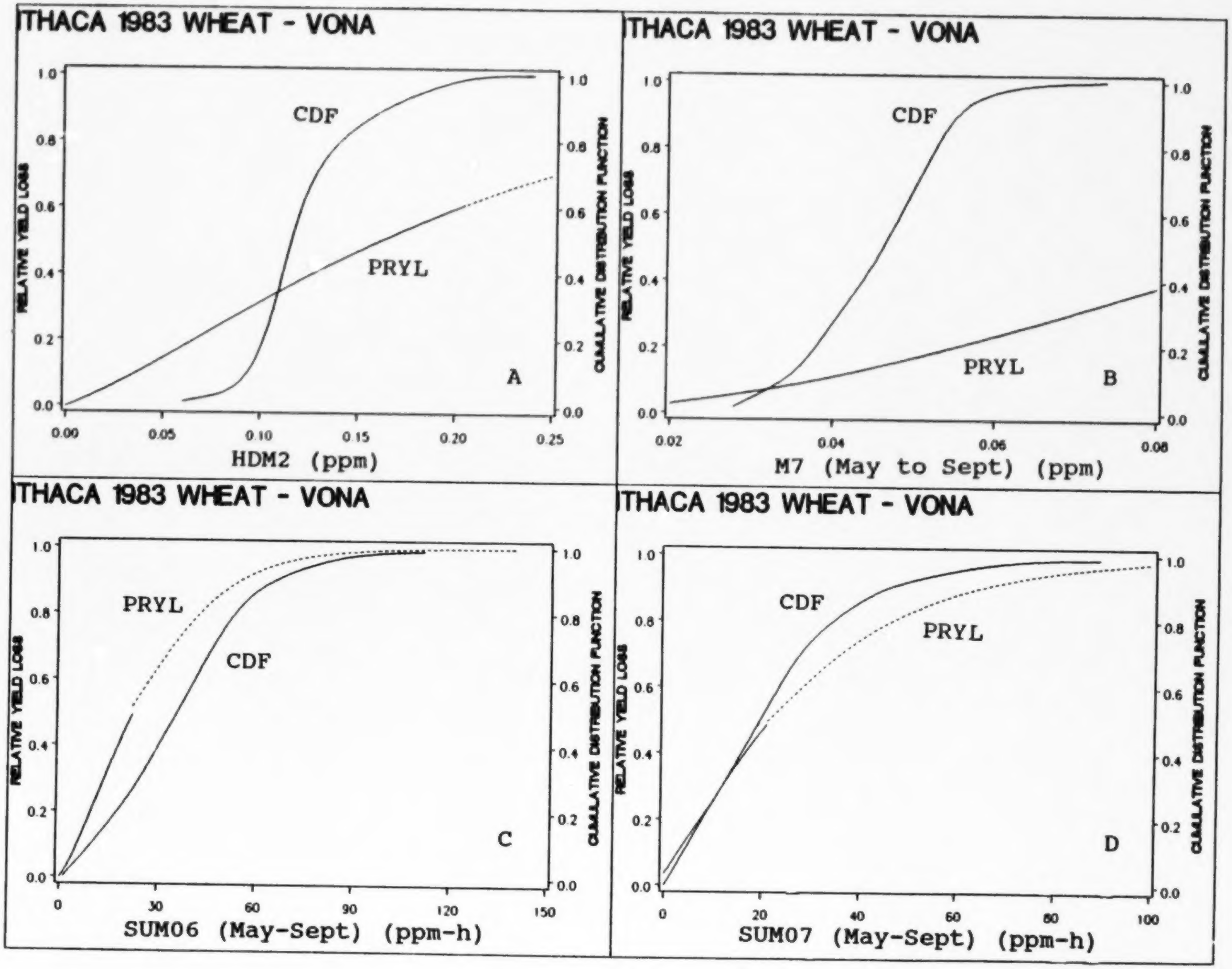


Figure B.11. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1983 wheat study at Ithaca, NY using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

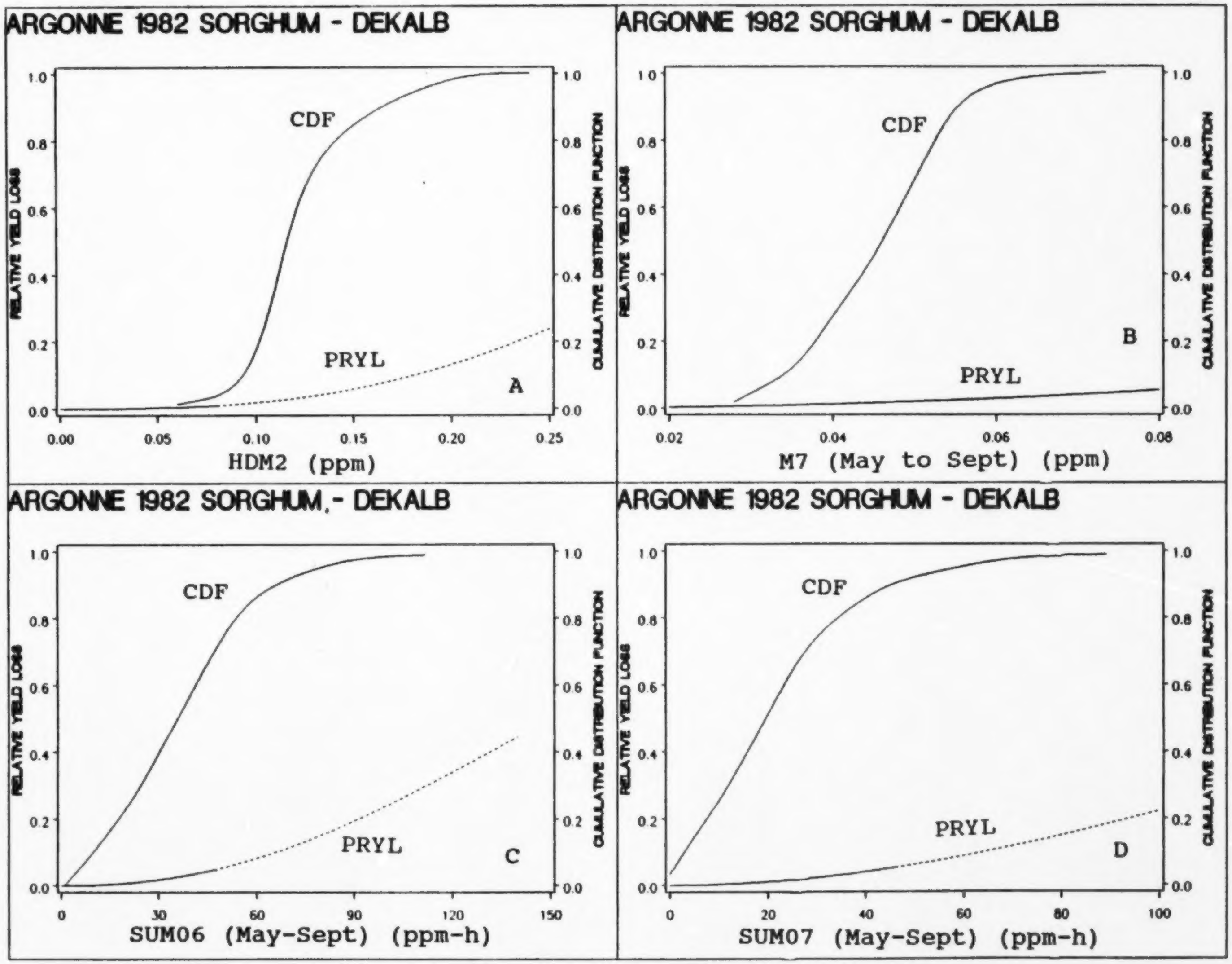


Figure B.12. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1982 sorghum study at Argonne, IL using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

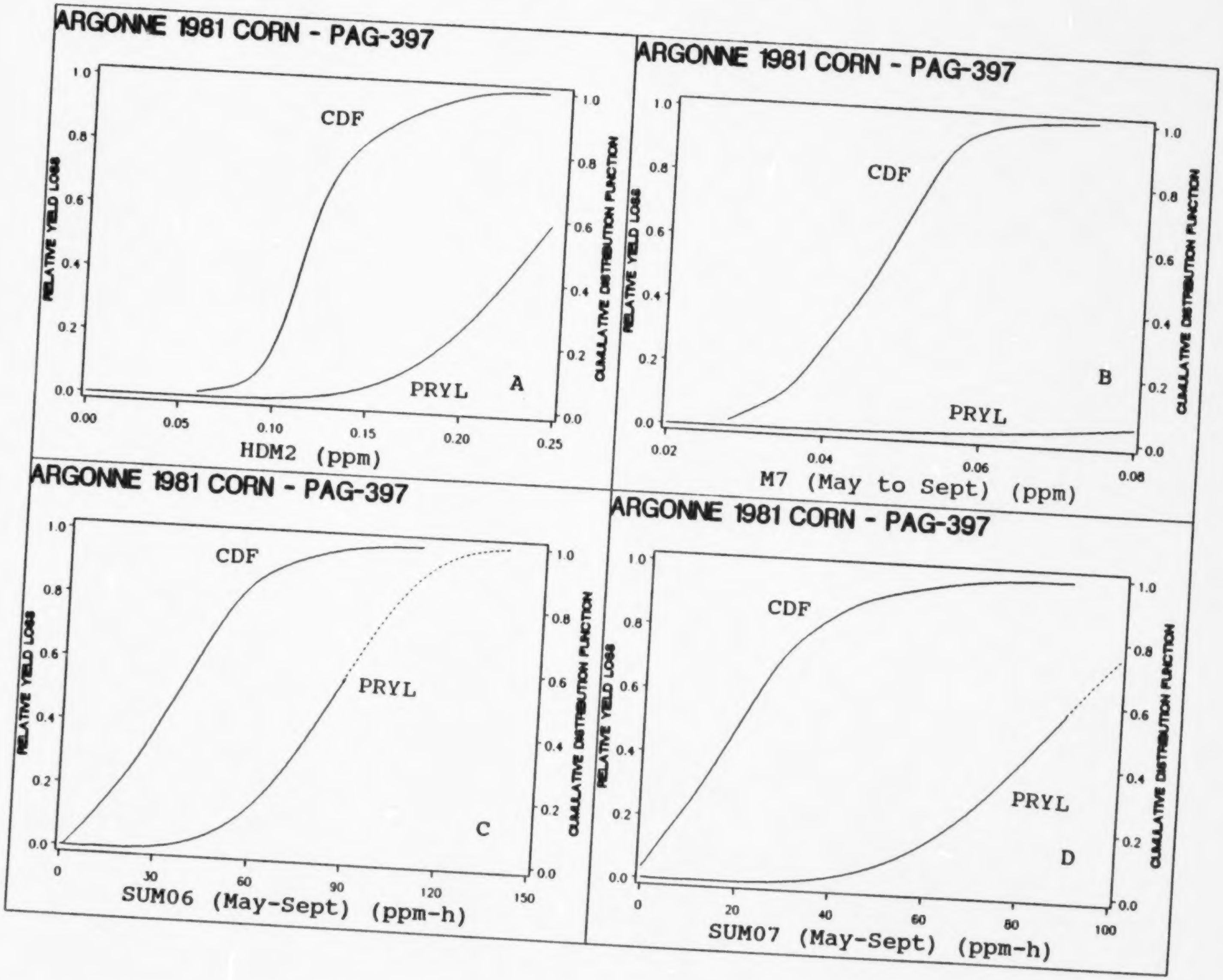


Figure B.13. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1981 corn study (Pag-397) at Argonne, IL using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

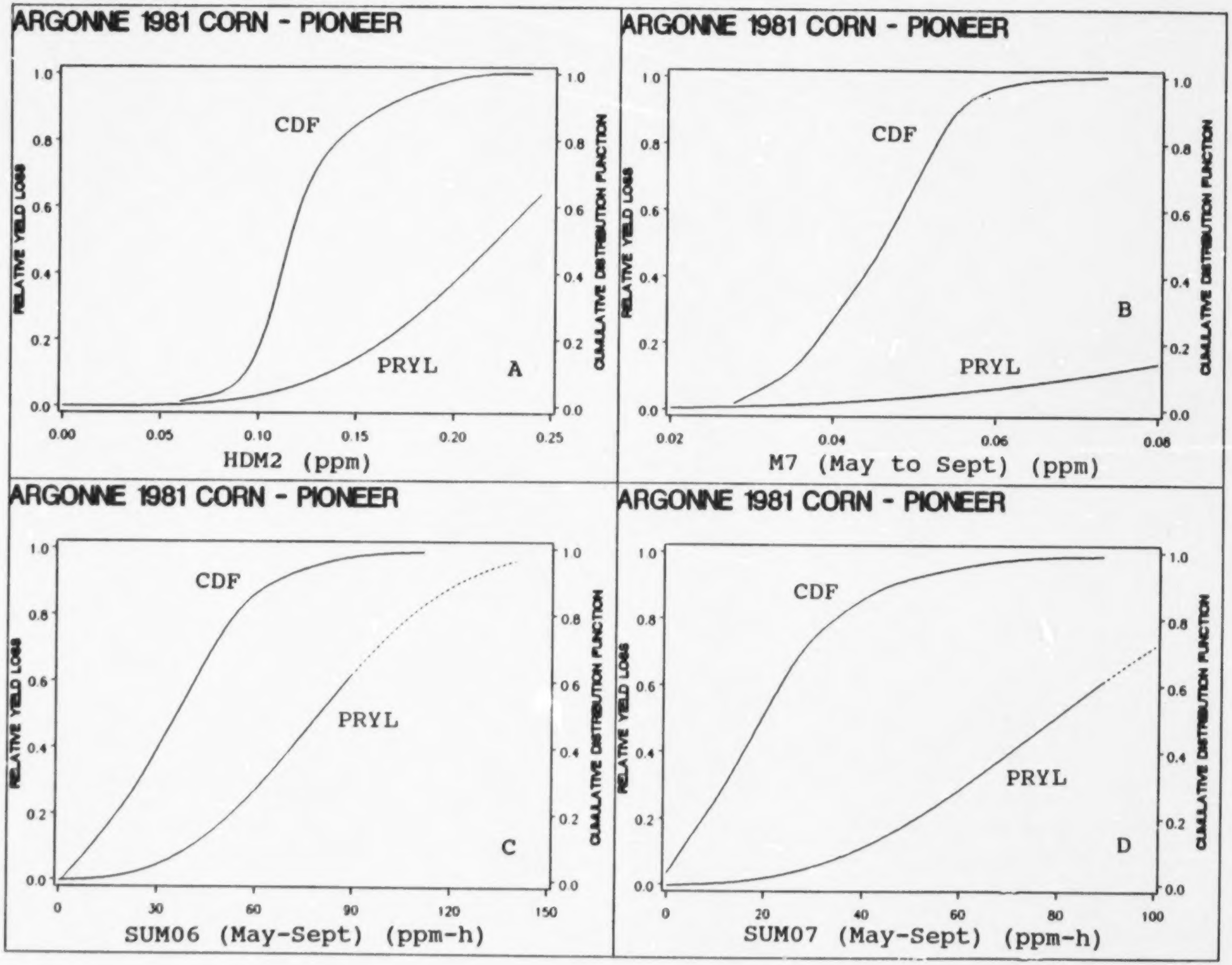


Figure B.14. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1981 corn study (Pioneer) at Argonne, IL using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

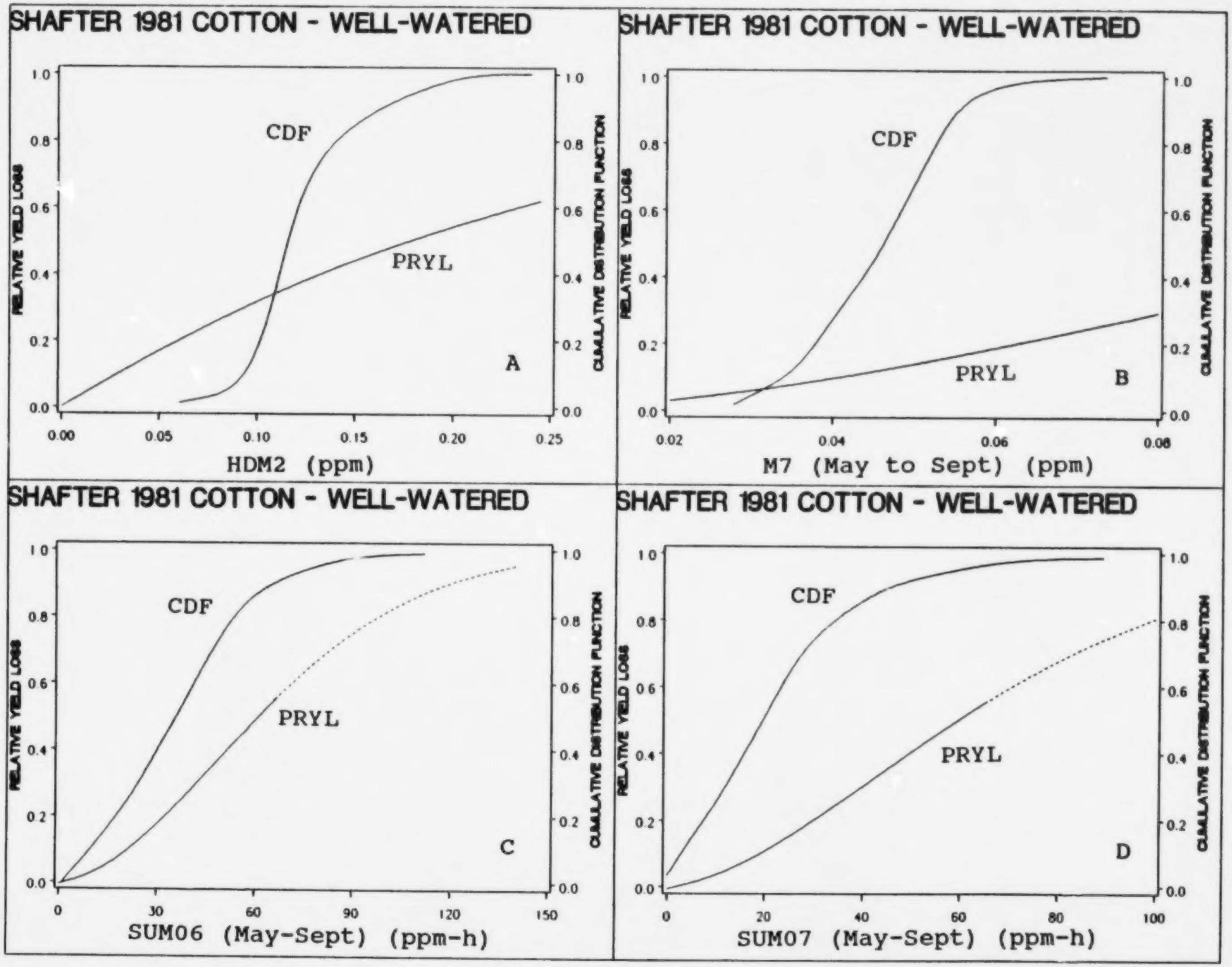


Figure B.15. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1981 cotton study at Shafter, CA using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

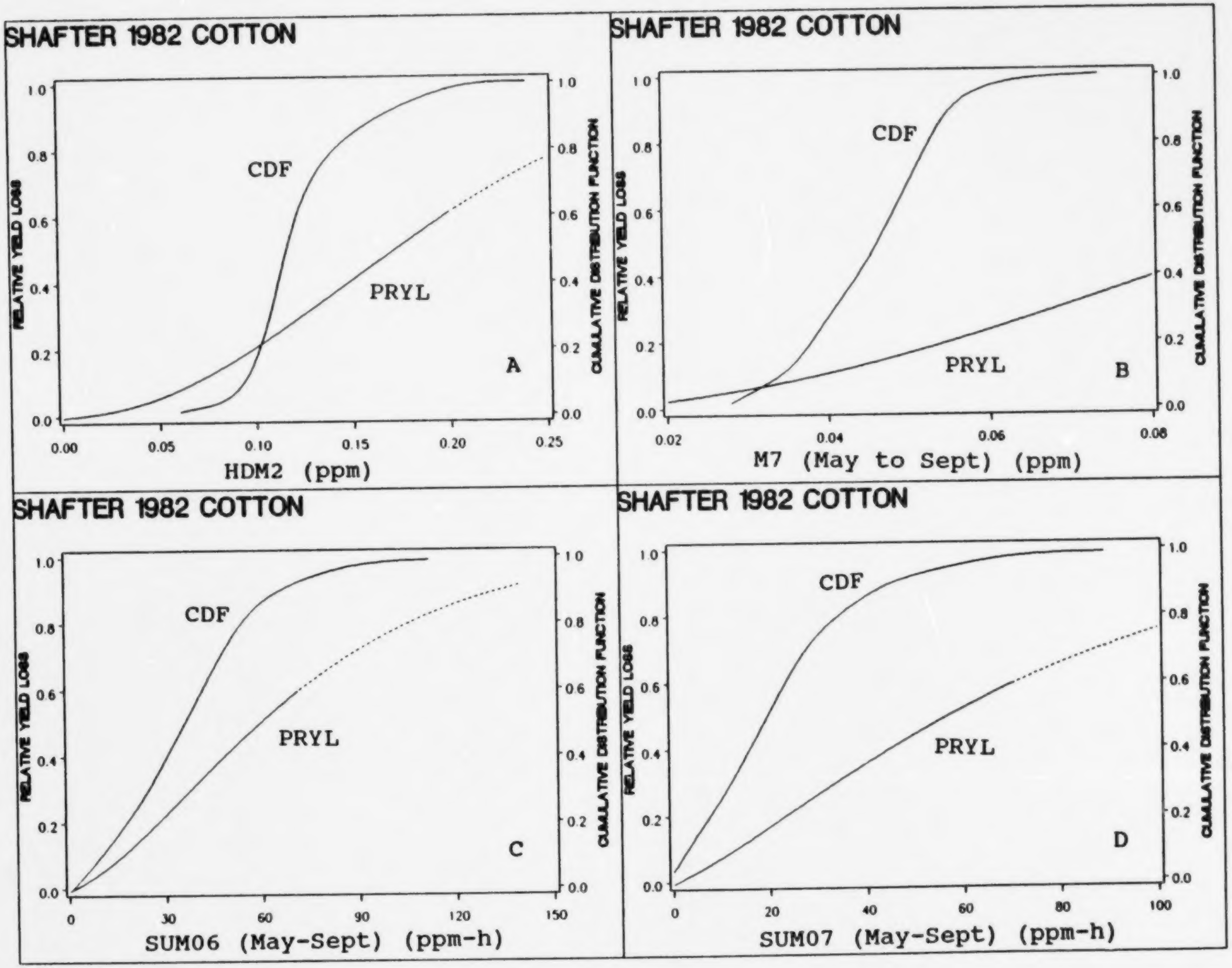


Figure B.16. Interrelationship between predicted relative yield loss (PRYL) functions estimated for the 1982 cotton study at Shafter, CA using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07, and corresponding cumulative distribution functions (CDFs) of exposure indices for the 83 agricultural site-years.

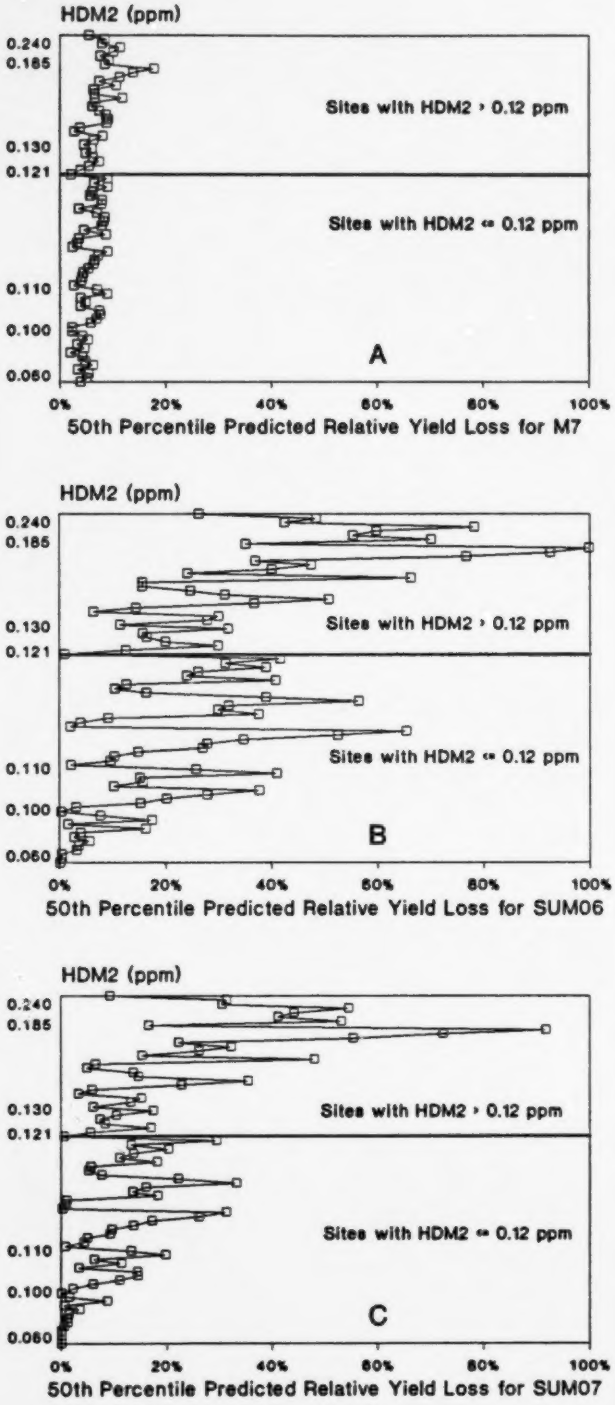


Figure B.17. Scatterplot of the 50th percentile of predicted relative yield losses (PRYLs) using (A) M7, (B) SUM06, and (C) SUM07, at the 83 agricultural site-years sorted in ascending order of HDM2.

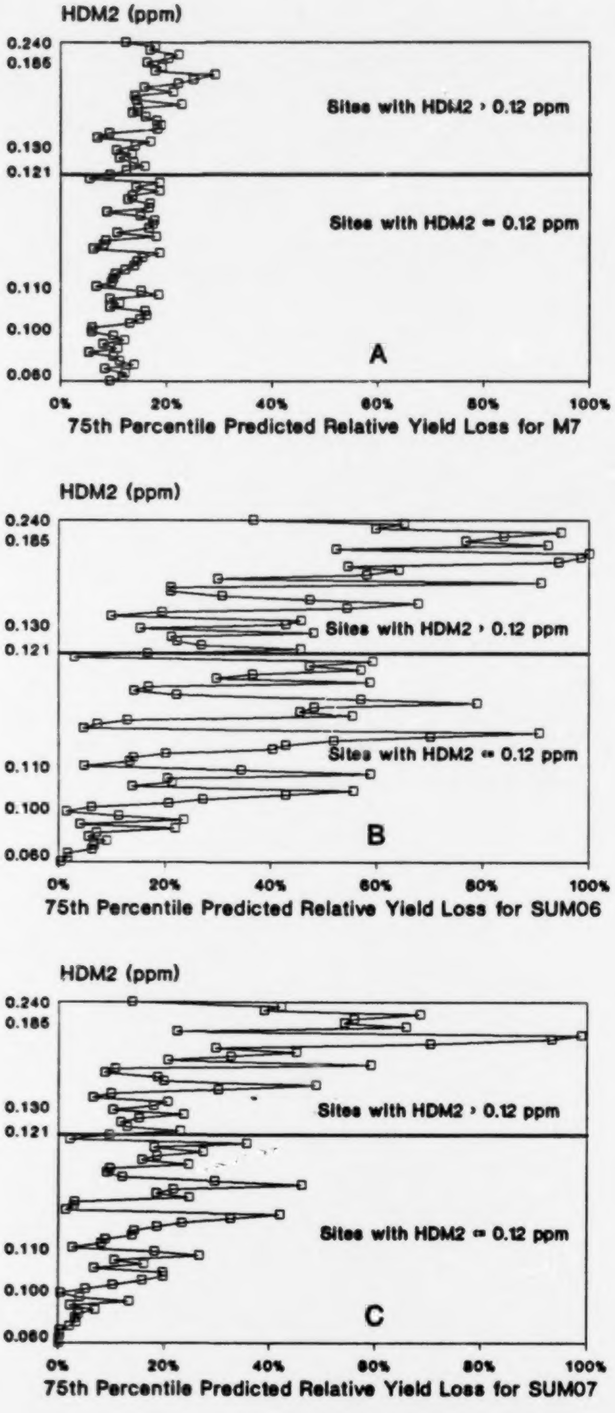


Figure B.18. Scatterplot of the 75th percentile of predicted relative yield losses (PRYLs) using (A) M7, (B) SUM06, and (C) SUM07, at the 83 agricultural site-years sorted in ascending order of HDM2.

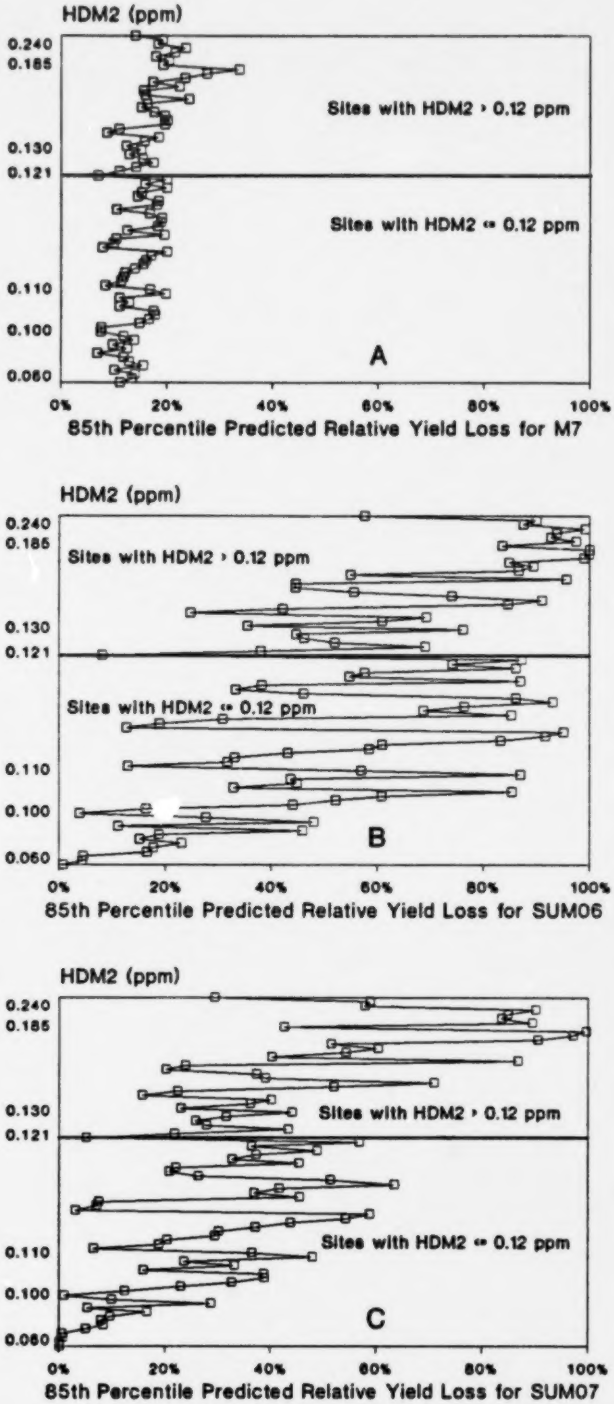


Figure B.19. Scatterplot of the 85th percentile of predicted relative yield losses (PRYLs) using (A) M7, (B) SUM06, and (C) SUM07, at the 83 agricultural site-years sorted in ascending order of HDM2.

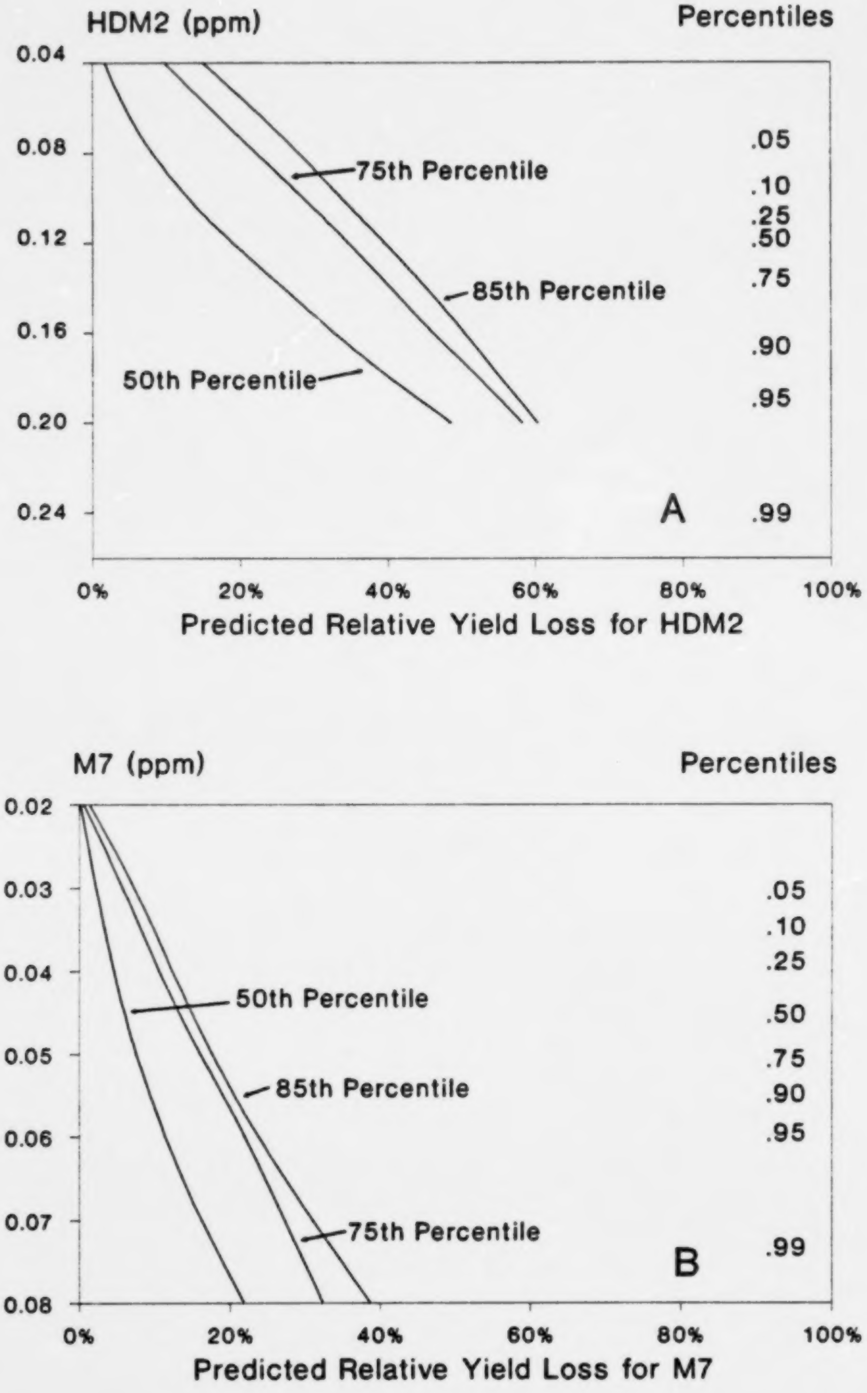


Figure B.20. The 50th, 75th, and 85th percentile of predicted relative yield losses for the 16 NCLAN cases using (A) HDM2, (B) M7, (C) SUM06, and (D) SUM07 calculated at 83 non-urban areas. Percentiles of the empirical distributions of air quality indicators at the 83 areas (Table 4.3 in Chapter 4) are given.

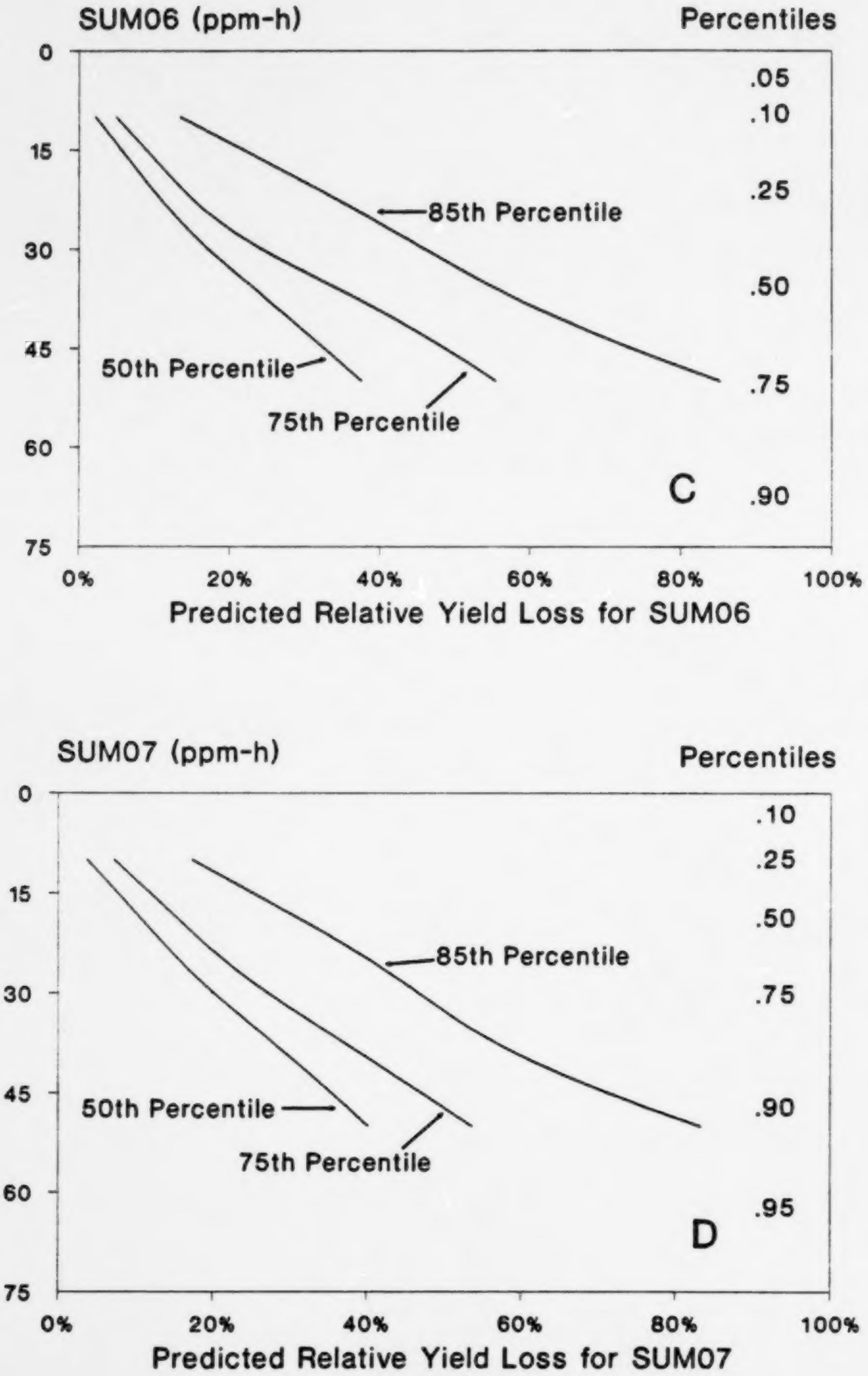


Figure B.20. (cont)